Characterization and Predictability of Low Flows based on Conceptual Modeling

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der Universität Zürich

von

Maria Staudinger

aus

Deutschland

Promotionskommitee

Prof. Dr. Jan Seibert (Vorsitz)
Dr. Kerstin Stahl

Zürich, 2015

Acknowledgments

This thesis would not have been possible without the financial support of the SNF which is gratefully acknowledged.

This thesis would as well not have been possible and all the years that led finally to it would not have been possibly as enjoyable without the help, support and distraction of many colleagues, my family and friends. It is not possible to thank you all adequately and individually here, but I appreciate each bit of it! I would like to thank some persons specifically in the following:

First of all I want to thank Jan Seibert, my main supervisor, which supported me during my PhD studies including two pregnancies and follow up periods. He had friendly words at times when it was least expected. Thanks also for the incredible patience with finalizing papers!

Thanks to Kerstin Stahl for supervising me during the PhD studies not only actively by revising articles and letters, but also for being an example in how to move and communicate in science.

Thanks to Markus Weiler for spontaneous solutions and inspiring ideas.

Thanks to my colleagues especially Michael Rinderer, Nans Addor and Benjamin Fischer for lending their ears and supporting me with sometimes fruitful discussions as well as many thanks to Tracy Ewen for checking the language of the articles over and over again.

Thanks to Klemens Rosin for first help in R in case of need.

Thanks to Andreas Hartmann and Michael Stölzle for revising parts of this thesis and more general discussions.

I also want to thank my sons Giulio and Linus that were keeping me away from working but also helped me grounding when it was necessary. Thanks also for not seriously destroying my computer or biting the power connections during the time of the thesis work!

Last but not least thanks to Edoardo Alberti for his patience and love also during harder times. Special thanks for all the mornings when you got up first an prepared the mocca!

Zusammenfassung

In der Vergangenheit hat sich gezeigt, dass verschiedene geographische Regionen sehr unterschiedlich auf meteorologische Trockenheiten reagieren. Ähnliche meteorologische Bedingungen führten teils zu stark unterschiedlich ausgeprägten hydrologischen Trockenheiten. Die zentrale Frage dieser Dissertation war daher, die Ursachen für diese unterschiedlichen Reaktionen zu untersuchen. Die Arbeit besteht aus vier wissenschaftlichen Aufsätzen, die sich jeweils verschiedenen Aspekten dieser übergreifenden Fragestellung widmen:

Aufsatz 1
Mit dem Ausblick auf häufigere und stärkere Trockenheitsereignisse bedarf es eines besseren Verständnisses der Reaktionen verschiedener Gebiete auf Trockenheiten. Anstatt detailliert auf Einzelprozesse einzugehen, wurde in dieser Studie ein Modellexperiment mit zwei Trocknungsszenarien durchgeführt. Dabei waren die mittlere Höhe der Einzugsgebiete, die Einzugsgebietsgrösse und -steigung mit der jeweiligen Trockenheitssensitivität korreliert. Höhergelegene Einzugsgebiete mit steileren Steigungen waren nach dieser Analyse weniger trockenheitssensitiv als niedriger gelegene Einzugsgebiete. Der simulierte Bodenfeuchtespeicher war signifikant mit der Grösse der Einzuggebiete korreliert, wobei kleinere Einzugsgebiete weniger trockenheitssensitiv waren als grössere. Es konnte jedoch keine deutliche Korrelation zwischen Trockenheitssensitivität und Hydrogeologie gefunden werden. Die Methodik dieser Studie lässt sich in anderen Regionen wiederholen. Eine Rangfolge der verschiedenen Einzugsgebiete nach deren Sensitivität kann helfen, um zu entscheiden welche Einzugsgebiete anfälliger für Trockenheiten sind.

Aufsatz 2
Die Möglichkeit, Abfluss und andere hydrologische Variablen vorherzusagen, ist verknüpft mit der Persistenz der Einzugsgebiete. Je grösser die Persistenz ist, desto wichtiger sind die Startbedingungen relativ zu den Wetterbedingungen während der Vorhersageperiode. In diesem Aufsatz wurde der Einfluss der Startbedingung der Simulation auf Vorhersagen mithilfe verschieder Modellexperimente analysiert. Die Höhe der Einzugsgebiete zeigte sich dabei als wichtigster Faktor für die Vorhersagbarkeit von Niedrigwasser im Vergleich zwischen verschiedenen Einzugsgebieten. Für die Vorhersagen von verschiedenen Jahren für einzelne Einzugsgebiete war die Situation zu Beginn der Simulationen entscheidend. Eine geringere Bodenfeuchte, weniger gefüllte Grundwasserspeicher zu Beginn der Simulationen oder eine grössere Schneeakkumulation führten zu längeren Persistenzen. Entgegen der intuitiven Erwartung von kleineren Speichern in höheren Lagen aufgrund flachgründigerer Böden, ergaben sich grössere Speicher aufgrund grösserer Grundwasserspeicher. Dies gibt Anlass, die Anfälligkeit von Gebirgseinzugsgebieten auf hydrologische Trockenheiten zu überdenken, besonders mit Blick auf Auswirkungen der Klimaveränderung.

Aufsatz 3
Der vielfach angewandte standardisierte Niederschlagsindex (engl. Standardized Precipitation Index, SPI) wurde in dieser Studie zu einem neuen Trockenheitsindex (Standarized

Melt and Rainfall Index, SMRI) weiterentwickelt, welcher die für viele Gebiete wichtigen Schneespeichereffekte berücksichtigt. Der SMRI basiert nur auf Temperatur- und Niederschlagsdaten. Ein Test an sieben Schweizerischen Einzugsgebieten zeigte, dass der SMRI Beginn und Ende hydrologischer Trockenheitsereignisse genauer angab als der SPI. Mit wachsendem Schneeeinfluss auf den Abfluss eines Einzugsgebietes, nahm auch der Nutzen des SMRI zu. Der SMRI ist für ein verbessertes Wassermanagement in schneebeeinflussten Regionen geeignet und ist eine einfache Alternative zu komplexeren Modellansätzen, die neben Niederschlags- und Temperaturdaten zusätzliche Informationen benötigen.

Aufsatz 4
Eine Überarbeitung von Modellkonzepten hinsichtlich Niedrigwasser bedarf eines klaren Verständnisses der strukturellen Defizite eines Modells. Das Modelliersystem FUSE (Framework for Understanding Structural Errors) ermöglichte es, den Einfluss der Modellstruktur auf die Modellgüte bei der Modellierung von Niedrigwasser zu analysieren. Grundannahme war hierbei, dass verschiedene Modellstrukturen ausschlaggebend für Unterschiede in der Modellgüte sind. Für die Beurteilung der Modelgüte wurden verschiedene Zielfunktionen verwendet, die den Unterschied zwischen beobachteten und simulierten Daten quantifizieren. Während verschiedene Modelstrukturen zu guten Modellergebnissen führten, konnten einige Modellstrukturen als deutlich weniger geeignet identifiziert werden. Ein Hauptergebnis dieser Studie war, dass es einen Unterschied in der Modellgüte für Sommer- und Winterniedrigwasser und Rezession gab. Tatsächlich zeigten sich besonders auffällig schlechte strukturellen Modellkombinationen jahreszeitenspezifisch. Die in dieser Studie verwendete Methode, eine systematische Analyse der Strukturen hydrologischer Modelle innerhalb von FUSE mit Zielfunktionen, die speziell Niedrigwasser- und Rezessionsverhalten bewerten, war vielversprechend. Ein weiteres Fazit dieser Studie war, für Niedrigwassermodellierung nicht nur eine einzelne Zielfunktion zu bauen, sondern mehrere verschiedene Zielfunktionen zu betrachten.

Alle in dieser Dissertation betrachteten Einzugsgebiete sind zu einem gewissen Grad schneebeeinflusst. Die Ergebnisse zeigten, wie wichtig es in diesen Gebieten ist, Schneeprozesse explizit miteinzubeziehen, um sinnvolle Trockenheitsindizes und Niedrigwassermodellierung zu erhalten. Weniger deutlich als Schnee war auch der Grundwasserspeicher ein wichtiger Faktor bei der Niedrigwassermodellierung. Eine interessante Schlussfolgerung dieser Arbeit war, dass in höheren Lagen grössere Speicherkapazitäten zu finden sein können, als intuitiv erwartet. In schneebeeinflussten Gebieten sind die Effekte von Wetter und Einzugsgebietseigenschaften schwierig zu trennen. Schnee fungiert als Speicher in der Akkumulationsphase, aber als Quelle während der Schmelzphase. In diesen Gebieten ist es daher wichtig dieser Doppelfunktion des Schnees Rechnung zu tragen.
In dieser Dissertation konnten Methoden entwickelt werden, die nützlich für die Analyse konzeptioneller Modelle sind. Hier lag der Fokus auf geeigneten Modellstrukturen sowie auf dem Einfluss von Startbedingungen auf die Vorhersage. Weiter wurden Methoden entwickelt, die ein besseres Trockenheitsmanagement ermöglichen könnten. Diese Methoden zielen auf die generelle Trockenheitssensitivität von Einzugsgebieten ab sowie speziell auf das Trockenheitsmanagement in schneebeeinflussten Regionen.

Summary

Historical meteorological drought events show that different regions react very differently and with varying degree of severity of consequent hydrological droughts. The overarching question is therefore why they show these different reactions, i.e. what are the main drivers. The thesis is comprised of four scientific studies that address different aspects of this question:

Study 1
With the perspective of increasing frequency and increasing severity of droughts, a better understanding of the reaction of different systems to droughts is needed - even in todays water rich countries. Instead of focusing on single processes in a catchment, a modeling experiment with two progressively drying scenarios was conducted. The results suggest that catchment elevation, size and slope are the main controls on the sensitivity of the catchments to drought. Higher elevation catchments and catchments with steeper slopes were found to be less sensitive to droughts than lower elevation catchments and catchments with less steep slopes. The soil moisture storage was significantly correlated to catchment size, where smaller catchments were less sensitive to droughts than larger catchments. However, no clear connection between drought sensitivity and hydrogeology could be found. A similar analysis to the one in this study could be repeated at any other location. A ranking for the different catchments could be a starting point to decide on which catchments are more vulnerable to droughts in a regional context.

Study 2
The predictability of streamflow and other hydrological variables is highly connected to persistence. Here, the influence of initial conditions on the prediction was analyzed using different model experiment setups. Mean catchment elevation was found to be the main source of predictability for low flow. Further, drier initial states of soil moisture and groundwater and more snow accumulation lead to longer persistence estimates. In contrast to an intuitive expectation from shallow soils in higher elevations, the results indicate larger groundwater storages in higher elevation catchments. This may motivate a reconsideration of the sensitivity of mountainous catchments to low flows in a changing climate.

Study 3
To account for snow storage effects and allow for regional comparisons, a new drought index, the Standarized Melt and Rainfall Index (SMRI), was developed. One advantage of the SMRI, is that it is based on temperature and precipitation data only. For seven Swiss catchments with increasing degree of snow influence, the SMRI was shown to be a good indicator for the onset and end of hydrological drought events as seen in streamflow. The more snow-influenced a catchment was, the better the SMRI described hydrological drought conditions as a complementary index to the SPI. The SMRI was tested in several Swiss catchments that showed some climatological differences even though they all are located in the temperate humid climate of Switzerland. Hence, the SMRI is a useful measure for water management in snow influenced regions and is a simple alternative to

a more complicated modeling approach that requires additional information to temperature and precipitation data.

Study 4
A revision of model concepts regarding low flows requires a clear understanding of the model's structural deficits. With the FUSE framework it was possible to look particularly at the influence of the model structure on the performance of low flow modeling. The basic assumption in the thesis was that different model structures are the reason for the differences in model performance. While most well performing models did not allow for the detection of a systematic influence of a model structure combination on the model performance, poor performance was more clearly linked to specific model structures. One main finding of this study was that there is a difference in model performance for summer and winter low flow and recession. In fact, all the structural decision combinations that were salient in this study were season specific. The method itself, i.e. a systematic analysis of the structures of hydrological models within the FUSE framework (Framework for Understanding Structural Errors), using objective functions targeting low flow and recession behavior, was promising. For low flow modeling it seems appropriate to use multiple objective functions and not to rely too much on a single function that is based on a comparison between simulated and observed data.

All catchments investigated in these studies are to a certain degree snow influenced. The results of the thesis showed the importance of the explicit consideration of snow processes in hydrological drought indices as well as in low flow modeling for these catchments. Less pronounced than snow storage, groundwater storage was also found to be important for studying and modeling low flow. One conclusion of the thesis was that larger storage capacities can be assumed for higher elevation catchments than intuitively expected.
In snow influenced catchments it is more difficult to distinguish between the influence of weather or catchment properties for low flow periods as snow acts as a water storage when it is accumulating and as a water source when it is melting. Especially in snow influenced catchments this effect has to be acknowledged, as shown in the thesis.
In this thesis, methods were developed, which are useful for analyzing conceptual models for low flow regarding their structure and the influence of initial conditions on the prediction. Further, an analysis to estimate the general drought vulnerability of catchments was developed as well as an index for the management of hydrological droughts, in particular, for snow influenced regions.

Contents

1

Introduction

1.1 Impacts of low flow

Droughts ultimately originate from a lack of precipitation and this precipitation deficit propagates through the hydrological cycle and affects various other components of the hydrological cycle (e.g., Eltahir and Yeh, 1999; Peters et al., 2003; Tallaksen and van Lanen, 2004; Van Loon and Van Lanen, 2011). On its way through the hydrological cycle the water deficit causes various impacts on nature and society. Few extreme events cause as much economic and ecologic damage as drought, which affects millions of people in the world each year (Wilhite, 2002).
One aspect of the hydrological cycle is surface water in lakes, reservoirs and rivers. From the full range of drought impacts many are related to surface water in particular. Associated low flows have a wide range of impacts including limited water supply to population centers, impacts on agriculture and forestry, waterborne transportation and power production (Stahl et al., 2012) (Table 1.1). Generally, more problematic in a low flow situation than in a medium or even high flow period is the contaminant dilution (Barnes and Kalita, 2001; Jordan et al., 1997). It is challenging to ensure safe concentrations of contaminants associated with wastewater (Smakhtin, 2001) and contaminants that enter the stream via soil or groundwater are most highly concentrated during low flow periods. During low flow situations the ecology within the stream is more vulnerable (e.g., Bunn and Arthington, 2002; Boulton, 2003; Bradford and Heinonen, 2008) as low flow leads to changed stream width, warmer streamflow temperatures, lower dissolved organic carbon and higher nutrient concentrations. Depending on the degree of reduced streamflow this can be very stressful for sensitive species (Price, 2011).

1.2 Definition of low flow

Often low flow and drought are mentioned in the same context sometimes without making a distinction between the two terms. Low flow is the "flow of water in a stream during prolonged dry weather" (WMO, 1974). Droughts are complex phenomena and thus there

Table 1.1: Impacts of hydrological droughts and low flows (after Stahl et al., 2012)

Impact category	Explanation
Agriculture and livestock farming	Reduction of cultivated areas due to a lack of irrigation water
	Regional shortage of feed/water for livestock
Aquaculture and fisheries	Reduced freshwater fishery production
	Reduced aquaculture production
Energy and industry	Reduced hydro power production
	Impaired production / shut down of thermal/ nuclear power plants (due to a lack of cooling water and/or environmental legislation for discharges into streams)
	Restriction / disruption of industrial production process (due to a lack of process water and/or environmental legislation/restrictions for discharges into streams)
Waterborne transportation	Impaired navigability of streams (reduction of load, increased need of interim storage of goods at ports)
	Stream closed for navigation
Tourism and recreation	Sport / recreation facilities affected by a lack of water
	Impaired use / navigability of surface waters for water sport activities
Water supply/ industries	Local water supply shortage/problems -drying up of local springs/wells
	Regional/region-wide water supply shortage/problems
	Drying up of reservoirs
	Limitations in water supply to households (including bans on domestic water use, supply cuts, need to ensure water supply by means of water transfers or bottled water)
Water quality	Increased temperature in surface waters
	Water quality deterioration/problems of surface waters; e.g. significant change of physio-chemical indicators, increased concentrations of pollutants, decreased oxygen saturation levels, eutrophication, algae bloom)
	Increased salinity of surface waters (saltwater intrusion and estuarine effects)
	Problems with groundwater quality
	Increased salinity of groundwater
	Problems with drinking water quality (e.g., increased treatment, violation of standards)
	Problems with bathing water quality
	Problems with irrigation water quality
	Problems with water quality for use in industrial production processes
Freshwater ecosystems: habitats, plants and wildlife	Increased mortality of aquatic species
	Migration and concentration (loss of wildlife in some areas and too many in others)
	Increased populations of invasive aquatic species
	Adverse impacts on populations of rare/endangered (protected) riparian species
	Loss of biodiversity
	Violation of environmental /minimum flow threshold
	Drying up of shallow water areas, weed growth or algae bloom
	Drying up of perennial stream sections
	Drying up of lakes and reservoirs (which have a habitat function)
	Mid-/long-term or even irreversible deterioration of wetlands

are various definitions available (Tallaksen and van Lanen, 2004; Mishra and Singh, 2010; Sheffield and Wood, 2012). In this thesis drought is used as defined by Tallaksen and van Lanen (2004) as “the below average natural water availability” and most of the times referring to streamflow. Smakhtin (2001) criticizes that the definition of low flow by WMO (1974), does not make a clear distinction between drought and low flow. He states that low flows are seasonal phenomena that are components of the flow regime of any river. Nevertheless, hydrological drought and low flow are intertwined: Low flow situations that are extended in their usual duration or that occur during uncommon seasons develop into hydrological droughts and hydrological droughts include low flow periods. However, a contiguous seasonal low flow situation does not necessarily develop to a drought (Smakhtin, 2001). In this thesis low flow shall be defined as any low streamflow that occurs may it be within the normal range of a certain flow regime or more extreme.

1.3 Low flow processes

Because of the connections between drought and low flow knowledge about low flow processes is a basis for understanding hydrological droughts. To understand low flow, numerous physiographic factors of a catchment need to be considered. The various aspects of the low flow regime of a stream include soil and aquifer properties, the rate, frequency and amount of recharge (Stoelzle et al., 2014b; Stoelzle et al., 2014a) as well as evapotranspiration from the catchment. Additionally, information about topography, vegetation and climate is needed (Smakhtin, 2001). Climate influence and catchment specific characteristics ultimately shape a low flow regime, its anomalies that may result in droughts and their re-occurrence frequency.

Some processes feed the low flow such as release from groundwater storage and soil storages, but also release from lakes and glaciers (Smakhtin, 2001). Glacier melt causes even higher flow during periods at times when catchments without glacier would show low flow. Other processes cause loss to low flow: transmission losses such as direct evaporation from standing or flowing water, evapotranspiration from seepage areas, groundwater recharge from streamflow, river bed losses, losses to relatively dry soils in the riparian zone (Smakhtin, 2001) as well as processes connected to ice and snow during winter. There are a number of human influences that affect low flow processes (Wang and Cai, 2009). Most of these anthropogenic influences intensify the natural processes (Smakhtin, 2001). However, in this thesis the focus is not on anthropogenic influences on low flow, but rather on natural processes connected to low flows and hydrological droughts.

Nevertheless, all the above mentioned factors and processes should be kept in mind as summer and winter low flows, for example, result from different hydrological processes. Summer low flows develop out of a deficit in precipitation and increased evapotranspiration. Early snow melt can also cause low flow in summer: Snow melt on still frozen topsoils may lead to extensive overland flow, which results in low soil moisture contents and thus probably reduced recharge (Tallaksen and van Lanen, 2004). This again controls groundwater discharge that appears in stream flow. Winter low flows are mainly due to storage of water in ice and snow and therefore the water cannot be found in the stream. Melt that occurs late in the season might lead to low flow in the following period because the groundwater did not receive recharge and the streams no groundwater

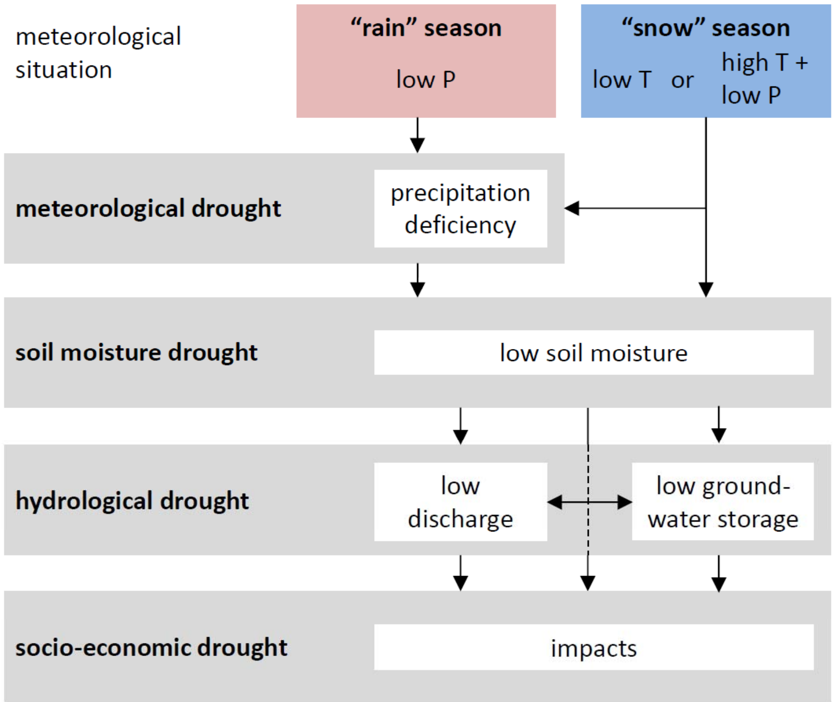

Figure 1.1: Propagation of drought (Van Loon, 2013).

discharge since the start of the winter.
From the low flow development it can be a very small step to the drought development and the drought propagation as it is described by Van Loon (2013) includes the above mentioned processes (Figure 1.1). A meteorological drought can develop to a hydrological drought in different ways (Eltahir and Yeh, 1999; Peters et al., 2003; Tallaksen and van Lanen, 2004; Van Loon and Van Lanen, 2011) that are controlled by catchment characteristics as well as climate. Several consequent meteorological droughts can turn into a combined and prolonged hydrological drought, they can be attenuated by the storages of a catchment and further there is often a varying time lag between meteorological drought, soil moisture drought and hydrological drought (Van Loon and Van Lanen, 2011; Haslinger et al., 2014). Apart of a deficit in precipitation hydrological droughts can also be caused by temporary storage of water in ice and snow (Van Loon et al., 2010). Based on this diversity of drought generating mechanisms only, Van Loon and Van Lanen (2012) developed a general hydrological drought typology and could distinguish between six different drought types that include the preceding type of precipitation and the preceding air temperature conditions (classical rainfall deficit drought, rain-to-snow-season drought, wet-to-dry-season drought, cold-snow-season drought, warm-snow-season drought, and composite drought).

1.4 Recent drought events in Europe

There were several noticeable drought events that affected Europe in the last decades. Important droughts of the earlier last century affecting Switzerland were summarized and their socio-economic impacts analyzed by Schorer (1992). The European Drought Reference (Stagge et al., 2013) informs about important droughts in Europe, their characteristics, climatological setting and their impacts. The last recent droughts that were affecting larger parts of Europe and also the catchments that were used in this thesis (Chapter 2) were in spring 2011 and in summer 2003. In the research articles of this thesis these two droughts are mentioned repeatedly and hence a short characterization of the summer drought in 2003 and of the spring drought in 2011 is given in the following.
The summer of 2003 brought extremely high air temperature conditions combined with less than average precipitation. The cause for this conditions was an anticyclone that was stationary over the Western part of Europe. This anticyclone brought hot air masses from North Africa to Europe and blocked cyclones on the Atlantic that could have brought precipitation to Europe (Rebetez et al., 2006). In Switzerland as in large areas of Europe all temperature records were beaten: The long-term average of air temperature was exceeded by about 3°C (Schär et al., 2004). The months February to September were additionally extremely dry: In many parts of Switzerland there was only half of the average precipitation that would normally fall in these months (BUWAL et al., 2004). Consequences to these meteorological conditions were extreme low flow conditions in many rivers in Europe. Some rivers also showed the opposite response, i.e. very high flows, as the high temperatures caused a lot of glacier melt. Depending on the region, in September to October it rained a lot so that the rivers were back to normal or even high flows in fall. In summary, the drought in summer 2003 was a drought driven by a lack of precipitation and high temperatures.
In spring 2011 there was only very little precipitation. The streamflow levels were extremely low already in spring since there was little snow in winter and thus little snow melt contributing to streamflow. The lowest flow occurred in May. Already in summer the low flow conditions were terminated by precipitation. From January to May in the central plain of Switzerland fell only 25 to 40% of the normal precipitation (Schlegel et al., 2011). Additionally March and April were over-averagely warm. Hence, beginning of May the streamflow was in parts of Switzerland so low as never before in this season of the year. The effect of the meteorological drought and high air temperatures was not only seen in streamflow, but also in groundwater levels (BAFU, 2011). A detailed description and analysis of the low flow situation in Germany in spring 2011 is given by (Kohn et al., 2014). The drought in spring 2011 can be summarized as a development of the preceding winter conditions followed by a lack of precipitation in combination with high temperatures during spring time.

1.5 Description of low flow and droughts

Various drought and low flow indices and metrics were developed such as the MAM-n-day (Tallaksen and van Lanen, 2004), spell duration (Zelenhasic and Salvai, 1987), deficit volume (Tallaksen and van Lanen, 2004) or base flow index (Institute of Hydrology, 1980). Along with those examples one of the most informative measures is the flow duration curve (FDC) which is not explicitly targeting low flow. However, it can be constructed targeting low flow by looking at seasonal flow values instead of using all the flow values of a year (Smakhtin, 2001). A number of low-flow indices may be estimated from the 75 - 95% time exceedance range of the FDC which is in most studies regarded as the range for low flow (Clausen and Biggs, 2000; Smakhtin, 2001; Burn et al., 2008). There exist various drought indices that make use of different variables; practically all use precipitation either alone or in combination with, for instance, temperature or soil moisture (Mishra and Singh, 2010). Some of the indices are based on local/regional information and cannot be easily compared to other regions as e.g. the surface water supply index (Shafer and Dezman, 1982; Doesken et al., 1991). Other indices like the Standardized Precipitation Index (SPI) are standardized and allow for comparability. Reviews of drought indices can be found in Heim Jr (2002), Keyantash and Dracup (2002), Tallaksen and van Lanen (2004), Mishra and Singh (2010), and Sheffield and Wood (2012). As drought indices include different parts of the hydrological cycle the appropriate selection of indices is important for impact studies. In a comparison study for the uncertainties connected to drought future projections, for instance, Burke and Brown (2008) found regions where different drought indices showed a common increase in drought connected to precipitation decrease, while other regions revealed that a change in drought was dependent on the definition of the drought index.

1.5.1 Standardized drought indices

The Standardized Precipitation Index (SPI) is an indicator for drought that was first introduced by McKee et al. (1993). Since its introduction, the SPI has been applied in many studies, in operational drought monitoring in the present, and also in scenario predictions of drought for climate change impact assessment (e.g. Ji and Peters, 2003; Ghosh and Mujumdar, 2007; Naresh Kumar et al., 2009; Orlowsky and Seneviratne, 2012; Naresh Kumar et al., 2009). A major advantage of the SPI compared to other drought indices is that it requires only precipitation data to describe drought severity. It is calculated based on a theoretical probability distribution fitted to the long-term precipitation record aggregated over a chosen preceding period. This probability distribution is then transformed into a normal distribution so that the mean SPI is zero. Positive SPI values indicate greater than mean precipitation, and negative values indicate less than mean precipitation. As the SPI is standardized, wetter and drier climates are represented in the same way allowing for regional comparison studies (Hayes et al., 1999).

Different precipitation aggregation periods can reflect the impact of drought as it propagates through the hydrological cycle into soil, streamflow and groundwater. Soil moisture conditions are related to precipitation anomalies on a relatively short scale, whereas streamflow, for instance, reflects longer-term precipitation anomalies (Hayes et al., 1999). While a hydrological drought ultimately results from a lack of precipitation, the time lag between the two may differ widely for different hydrological systems. SPI and other drought indices based on climatic variables were often used as indicators for agricultural drought. However, with the right aggregation time a climatic drought index such as the SPI may also be a suitable indicator for hydrological drought. Hydrological drought occurs with a time lag and in order to use an SPI as an indicator for hydrological drought, the relevant time lag needs to be known. The US Drought Monitor, for example, uses composite drought indices with a focus on short SPI aggregation periods for warnings on agricultural drought impacts and composite indices with a focus on longer SPI aggregation periods for warnings on hydrological drought impacts (`droughtmonitor.unl.edu`). Several studies have investigated this time lag in order to find the most suitable SPI aggregation period for hydrological drought characterization. They have found that due to highly variable catchment storage processes that affect the propagation of drought, this time lag is difficult to determine.

Shukla and Wood (2008) found the differences between the hydrological and climatological drought indices increasing with decreasing lag times (12-, 6-, 3-, 1-month lag). Vidal et al. (2010) generated space-time fields of drought index values for different drought types and time scales in France, and found that the ranking of drought events is highly dependent on both the time scale and the driver considered. Haslinger et al. (2014) found the best time lag to be four months and relations between meteorological drought and streamflow drought in Austria to be significant, except for catchments with important groundwater storage and snow processes. Snow storage was also mentioned by Shukla and Wood (2008) as a reason for a missing link between meteorological drought and streamflow drought particularly for short lag times.

To create a methodologically consistent indicator of hydrological droughts, several studies have transferred the SPI approach to observed and modeled hydrological variables (Table 1.2). The SPI can be modified to move from a precipitation drought characterization to a hydrological drought characterization without the full complexity of a hydrological model by accounting for a known first-order control on catchment hydrology that affects drought. Recently, for instance, Vicente-Serrano et al. (2010) introduced an index that accounts for evapotranspiration as an important amplifier of drought.

1.5.2 Percentiles

A more general approach to analyze low flow periods is the use of streamflow percentiles. The FDC illustrates the frequency of streamflow as function of the percentage of time that the streamflow is exceeded (Vogel and Fennessey, 1994). Usually, the full observation period serves as basis for the construction of the FDC, but alternatively also only particular periods can be taken for the analysis (Smakhtin, 2001; Tallaksen and van Lanen, 2004). Expressing flows as percentiles allows for comparison of flow conditions in different catchments given that the FDC are normalized (Gustard et al., 1992). In connection to low flow and drought, empirical percentiles have been used mostly in studies that extract further drought characteristics below a threshold to define severity-area-duration or frequencies (and return periods) (e.g., Cancelliere and Salas, 2010; van Vliet et al., 2012; Gudmundsson et al., 2012). Smakhtin (2001) indicated that the commonly used low flow range of a flow duration curve is the 70% to 99% percentile range, or the Q_{70} to Q_{99} range. The Q_{95} and Q_{90} flows are most often used as low flow indices in the academic literature (Pyrce, 2004). The choice of a different percentile in the calculation of a threshold level for the extraction of further drought characteristics changes their magnitude. For example, with Q_{95} as threshold fewer events that last shorter are identified as compared to Q_{70} as threshold. However, the relation between the drought characteristics of various hydrometeorological variables and catchments is not expected to change (Van Loon, 2013; Ko and Tarhule, 1994; Tate and Freeman, 2000).

1.5.3 Recession analysis

The falling limb of a hydrograph, the hydrograph recession, is a catchment's response to a dry spell and it comprises the processes that occur during a low flow period (Tallaksen and van Lanen, 2004): the shape of the recession curve reflects the gradual depletion of water stored in a catchment during periods with little or no precipitation contributing to streamflow. Initially, the recession curve is steep as overland flow and/or interflow in a topsoil leave the basin. It flattens out with delayed water from deeper subsurface stores and may eventually become nearly constant if sustained by outflow from groundwater storage or from a glacier (Smakhtin, 2001). The recession curve in its flatter part is also called baseflow recession and it describes how different factors in a catchment influence the generation of flow in dry weather periods (Tallaksen and van Lanen, 2004). The most important catchment characteristics found to affect the recession rate are hydrogeology, relief and climate (Tallaksen, 1995). Catchments with a slow recession rate are typically groundwater dominated while impermeable catchments with little storage show a much faster recession behavior. A comparison of seasonal recession rates usually results in faster rates in summer than in autumn or winter (Tallaksen and van Lanen, 2004). Recession behavior can be quantified with the help of recession analysis (e.g., Sujono et al., 2004). In the long history of recession analysis, there have been many attempts to relate the recession behavior of rivers to the drainage from the catchment (see reviews by Hall, 1968; Tallaksen, 1995; Smakhtin, 2001) (as well as more recent studies Lamb, Beven, et al., 1997; Szilagyi et al., 1998; Wittenberg and Sivapalan, 1999; Wittenberg, 1999; Wittenberg, 2003; Szilagyi et al., 2007; Rupp and Woods, 2008; Kirchner, 2009). These attempts are based on the assumption that streamflow in rainless periods stems solely from stored water in a catchment and that hence, the recession curves should be characteristic for each specific catchment. If this assumption holds, recession analysis can also serve as a tool to improve low flow modeling. However, Stoelzle et al. (2013) concluded from their study using three widely used recession analysis methods and perturbations amongst those methods, that recession characteristics show limited comparability due to the distinctiveness of individual analysis methods.

1.6 Modeling low flow and droughts

For different management issues different models were designed and applied to help answering questions regarding frequency, return periods and drought intensities both in the present climate but also in the light of climate change as summarized by Mishra and Singh (2010). There are many studies that used scenarios to estimate the impact of climate change on streamflow in general and also some that focus on drought in particular (e.g. Wetherald and Manabe, 1999; Wetherald and Manabe, 2002; Wang, 2005; Lehner et al., 2006). The usual approach is to use simulations of general circulation models or regional climate models (GCM/RCM) with plausible scenarios of greenhouse gas emissions to drive hydrological models. However, there are large uncertainties connected to the GCM and RCM simulations and the choice of bias correction method (Teutschbein and Seibert, 2012; Teutschbein and Seibert, 2013). The range of resulting impacts is accordingly high. For low flow modeling usually hydrological models are applied. Hydrological models include simple statistical models with few parameters, conceptual models with varying complexity as well as even more complex physically-based models (e.g. Wagener and Wheater, 2004; Matonse and Kroll, 2009; Beven, 2011; Solomatine and Wagener, 2011). Conceptual models are simplified descriptions of hydrological processes. These models are built up by storage elements, that are filled by fluxes such as rainfall, infiltration and percolation and emptied by fluxes such as evapotranspiration, drainage and runoff. Parameters of conceptual models describe usually aggregated processes, which has the consequence that the model parameters are usually not directly measurable but are derived through calibration (Solomatine and Wagener, 2011). Conceptual models differ among each other as they are a priori specified by the modelers' concept and thus also the complexity in the model structures describing catchment hydrology varies (see e.g. Beven and Kirkby, 1979; Liang et al., 1994; Lindström et al., 1997a; Clark et al., 2008).

Table 1.2: Main drought indices.

Index	Variables	Description	Advantage	Disadvantage	Key publication
Palmer's Index	Precipitation, temperature	Precipitation and temperature analyzed in a water balance model; comparison of meteorological and hydrological drought across space and time	Comprehensive index used for the entire U.S.	Palmer values may lag droughts by several months; not suited for mountain areas or areas of frequent climatic extremes; complex, has an unspecified, built-in time scale that can be misleading	Palmer (1965)
Standardized Precipitation Index (SPI)	Precipitation	Precipitation as cumulative deficiency measure for meteorological drought; based on the probability of precipitation for any time scale	Can be computed for different time scales, can provide early warning of drought and help assess drought severity; is simple in both application and data requirement	Considers only precipitation	McKee et al. (1993)
Standardized Runoff Index (SRI)	Modelled runoff	Runoff index calculated based on the SPI methodology	Can complement the SPI especially in periods when variables other than precipitation become more important, e.g. periods of snow accumulation and melt	Requires a lot of data; grid runoff cannot be validated.	Shukla and Wood (2008)
Standardized Streamflow Index (SFI)	Observed Streamflow	Streamflow index calculated with the SPI methodology	Indicates hydrological droughts for any gauged catchment	Cannot be used in ungauged catchments	Vidal et al. (2010)
Standardized Precipitation and Evaporation Index ($SPEI$)	Precipitation, evaporation	Water balance index calculated with the SPI methodology	Considers evapotranspiration as key process	Uncertainty related to evapotranspiration computations; validation	Vicente-Serrano et al. (2010)
Surface Water Supply Index ($SWSI$)	Snow pack, reservoir storage, streamflow, and precipitation	Designed to complement the Palmer index in areas, where snow is a key element of water supply; calculated by river basin	Represents water supply conditions unique to each basin.	Changing a data collection, station or water management requires new algorithms; index is unique to each basin, which limits comparisons between catchments	Shafer and Dezman (1982) and Doesken et al. (1991)

Physically based models differ in the detail of the presentation of the processes and are based on the physical laws of conversation of mass, momentum and energy. They were developed with the intention that a detailed description would make the need of calibrated parameters unnecessary. However, since until today the problem of scale (measurement scale is not equal to simulation scale) remained with the consequence that not all parameters can be inferred directly from measurements, i.e. calibration is still required (Solomatine and Wagener, 2011). Additionally to the scale problem, this kind of models suffers a high demand of data and from over-parametrization (Beven, 1989). Reed et al. (2004) showed in a modeling comparison experiment that due to the difficulties in calibrating distributed models, lumped conceptual models are in many cases in fact more accurate in reproducing streamflow than distributed physical ones.

Traditionally rainfall runoff models are designed and calibrated to simulate peak flows as good as possible, as the first focus of those models was on floods. The application of these models specifically to low-flow situations, however, has been limited (Smakhtin, 2001). Low flow studies focused rather on statistical methods, such as indices and extreme value analysis (Gustard et al., 2008). During the process of modeling in general but also low flows in particular there are uncertainties arising from different sources (Melching, 1995; Gupta et al., 2005), which are related to the understanding and measurement capabilities (Solomatine and Wagener, 2011):

Perceptual uncertainty originates from the perceptual representation of the catchment that is subsequently translated into mathematical form of the model (Beven, 2011). This understanding might be poor especially with regard to subsurface processes (Solomatine and Wagener, 2011). Different hydrological processes may overlap and the specific reasons for low streamflow can not always be identified properly, which can hamper analysis and modeling of low flow and drought (Smakhtin, 2001; Gottschalk et al., 1997; Tallaksen and van Lanen, 2004).

Data uncertainty is caused by errors and imprecision in the measurement data, or by data processing. Regarding low flow, the lower the streamflow, the more difficult it is to measure precisely. It becomes even more difficult when freezing processes occur during winter time: hydrologists that work in northern catchments and are involved with those problems have to apply models to improve the quality of their measured data to overcome the changed stage discharge relation under ice cover conditions (Moore et al., 2002). Variations in the data are partly due to measurement errors and thus some authors suggest using a pre-treatment of the data. Reusser et al. (2009) for instance smooth their data with a moving average filter to crystallize the information and minimize the errors before they use their data.

Parameter estimation uncertainty arises if no unique best parameter set based on the available data can be located. This will commonly result in significant prediction uncertainty if the model is extrapolated to predict the system behavior under changed conditions or in similar systems (Solomatine and Wagener, 2011). This is critical also for low flow and drought predictions.

Model structural uncertainty is introduced because of simplifications, inadequacies and ambiguities in the description of the system (Solomatine and Wagener, 2011). A revision of model concepts regarding low flows requires a clear understanding of the model's structural deficits. A common approach to investigate the impact of the differences in model structure is to perform model inter-comparison experiments, e.g. Henderson-

Sellers et al. (1993), Reed et al. (2004), Duan et al. (2006), Breuer et al. (2009) and Holländer et al. (2009). Such experiments have been helpful to explore model simulation performance of lumped (Duan et al., 2006; Breuer et al., 2009), semi-distributed (Duan et al., 2006; Holländer et al., 2009) and distributed (Henderson-Sellers et al., 1993; Reed et al., 2004; Holländer et al., 2009) models in a consistent way using the same input data. The reasons for the differences, however, remain unclear since each model uses different interacting parameterizations to simulate the hydrological processes (Clark et al., 2008). Discrepancies between observed and simulated streamflow can arise from errors in the input data rather than weaknesses in model structure. This complicates the investigation of the impact of the differences in model structure. Clark et al. (2008) created a computational framework that enables a separate evaluation of each model component. The Framework for Understanding Structural Errors (FUSE) differs from others as it modularizes individual flux equations instead of linking available sub-models. The FUSE approach can help to get a better understanding of the hydrological processes occurring. Clark et al. (2008) first introduced FUSE, as a diagnostic tool to evaluate the performance of hydrological model structures for two climatically different catchments. However, no study so far was evaluating the representation of low flow processes by different model structures.

Numerical errors of fixed-step explicit schemes commonly used in hydrology can dwarf the structural errors of the model conceptualization (Clark and Kavetski, 2010). This does not necessarily degrade model predictions, but can even generate adequate performance for parameter sets where numerical errors compensate for structural errors. Numerical time stepping schemes are often still believed to be of little significance compared to the uncertainties in the data and governing equations. However, errors occurring from numerical time stepping were shown to be large and may even surpass conceptualization errors (Clark and Kavetski, 2010). Kavetski and Kuczera (2007) found further that problems believed to be endemic to environmental models were merely artifacts of the numerical implementation. Using an inappropriate time stepping scheme is a major methodological weakness as modeling errors from both physical conceptualization and numerical implementation become difficult to separate, diagnose and resolve (Clark and Kavetski, 2010).

Different sources of uncertainty may vary with model complexity (Figure 1.2). As the model complexity increases, model structural uncertainty decreases. With increasing complexity of the model, the number of inputs and parameters also increases and thus input and parameter uncertainty may increase. As of the trade off between model structure uncertainty and data/parameter uncertainty, every model has its optimal level of model complexity where the total uncertainty is minimal. (Solomatine and Wagener, 2011)

Due to to the limitations of model structures and and data most models need to be calibrated (Solomatine and Wagener, 2011; Beven, 2011) and there are only very few cases where model parameters were measured or estimated a priori (Beven, 1984; Parkin et al., 1996; Bathurst et al., 2004; Beven, 2011). Generally, the target of model calibration is to obtain a model having a consistent input-state-output behavior with the measurements of catchment behavior as well as accurate and precise predictions. Furthermore, the models structure and behavior are supposed to be consistent with current hydrologic understanding of reality (Gupta et al., 2005). Hydrological models can be

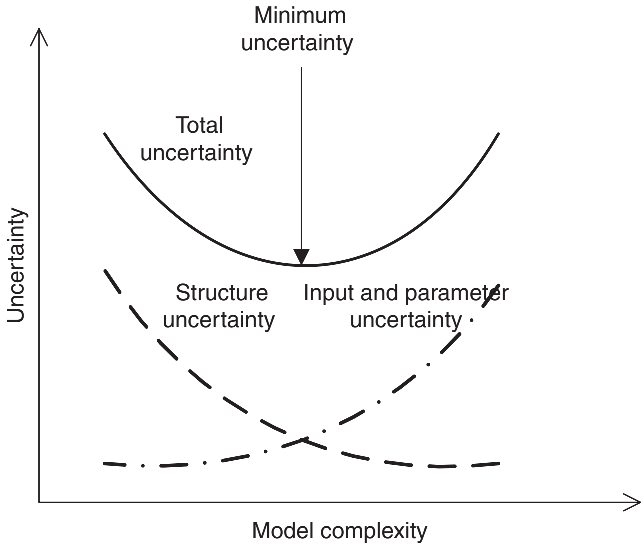

Figure 1.2: Various sources of uncertainty depending on the model complexity (Solomatine and Wagener, 2011).

calibrated manually or by using automatic optimization techniques and examples to different calibration algorithms can be found in e.g., Press et al. (1992), Duan et al. (1994), Sorooshian and Gupta (1995), Seibert (1999), and Beven (2011).
However, this thesis shall not give detailed descriptions of different calibration algorithms, but rather focus more on objective functions and their influence on the calibration results especially for low flow: model parameters are calibrated by optimization procedures, that require optimization criteria called objective functions. Dawson et al. (2007) classify objective functions commonly used in hydrology into three major groups 1) dimensionless coefficients, 2) relative error measures and 3) absolute error measures. The objective functions presented in Table 1.3 are organized according to this classification and the equations are throughout given with streamflow (Q) as example. Statistical measures that are implemented in many objective functions are based on Pearson's product-moment correlation coefficient r (Pearson, 1896), which describes the degree of collinearity between observed and simulated variates. However, r is limited as it standardizes the observed and predicted means and variances. This makes correlation based metrics insensitive to additive and proportional differences and thus good results will be indicated as long as the errors occur systematically (Dawson et al., 2007). The equations of correlation based metrics are based on a consideration of linear relationships and are not sufficient for rarely linear hydrological models. Further, correlation-based metrics are more sensitive to outliers than to observations near the mean (Moore and Notz, 2005). These metrics are still customary to determine how well a model simulates the observed data although they provide only a biased view of the efficiency of a model (Legates and McCabe Jr, 1999).
The Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970) is still among the most used objective functions in hydrological modeling. It is a normalized measure that com-

pares the mean squared error generated by a particular model simulation to the variance of the output. The NSE represents an improvement over simply correlation-based metrics for model evaluation as it is sensitive to differences in the observed and simulated means and variances. However, the NSE does not provide a reliable basis for comparing the results of different case studies: Schaefli and Gupta (2007) show that models that are intended to simulate seasonal variations can have a high NSE even though they poorly simulate variations on the smaller daily scale. Legates and McCabe Jr (1999) and Krause et al. (2005) stress the fact that the NSE in addition emphasizes high values of streamflow relative to other measurements because the deviations are squared, which makes it a poor objective function for low flow simulations.
Relative error metrics used for calibration set the differences between simulated and observed values in relation to the observed values, which makes them to metrics that do not favor extreme values. Considering conventional error metrics Teegavarapu and Elshorbagy (2005) conclude that they are of limited use and do not always allow a comprehensive assessment of the performance of the model developed for a specific application. Seibert (2001) and Schaefli and Gupta (2007) called for the use of benchmark models. For clarity reasons, it seems necessary to establish appropriate benchmark models - models having an explanatory power that is easy to apprehend for a given case study and a given modeling time step (Seibert, 2001). The definition of an appropriate benchmark model is particularly important when comparing model performance over a variety of hydrologic regimes (Schaefli and Gupta, 2007). The performance improvement of the hydrologic model over the benchmark model can be measured by defining a normalized benchmark efficiency. Examples of such benchmark models that are easy to apprehend are given by Schaefli and Gupta (2007). Another way to avoid a strong influence of the disadvantages connected the NSE and mean squared error was proposed by Gupta et al. (2009), which presented an objective function that facilitates the analysis of the relative importance of correlation, bias and variability.

Most of the objective functions used in hydrological modeling focus on how well the simulated hydrograph shape, flood peaks and flow volumes match with corresponding observed features. These criteria provide relatively little information about the quality of low streamflow simulations and it is necessary to consider objective functions, which reflect the model performance for low streamflow. This can be derived by transforming the variables in any objective function using logarithmic or Box-Cox (Box and Cox, 1964) transformations (Tallaksen and van Lanen, 2004). Some of the aforementioned objective functions explicitly emphasize peak flows and those functions do not seem to be appropriate as an objective function for low flow even after transformation. A review and evaluation of efficiency criteria suitable for evaluating low-flow simulations based on French catchments is given by Pushpalatha et al. (2012).

1.6.1 Drought early recognition

The basic objective of drought early recognition is to provide timely warning, so that damages can be reduced or even avoided. As already mentioned the severity of a drought depends clearly on the climatological deficit of water, but also on the hydrological system that is confronted with this deficit. To recognize locally critical conditions early and pro-

Table 1.3: Common objective functions in hydrology.

Class	Objective Function	Abbreviation	Range	Bias
Dimensionless coefficients	Coefficient of determination	R^2	[0,1]	to high values
	Coefficient of efficiency	NSE	[-inf,1]	to high values
	Index of agreement	d	[0,1]	to high values
	Persistence index	PI	[-inf,1]	to high values
Absolute errors	Absolute maximum error	AME	[0, inf]	-
	Peak difference	PDIFF	[0, inf]	-
	Mean absolute error	MAE	[0, inf]	-
	Mean error	ME	[0, inf]	to high values
	Root mean square error	RSME	[0, inf]	to high values
	Akaike Information Criterion	AIC	[0, inf]	-
	Bayesian Information Criterion	BIC	[0, inf]	-
Relative errors	relative absolute error	RAE	[0, inf]	scale dependent
	mean absolute relative error	MARE	[0, inf]	to low values
	median absolute percentage error	MdAPE	[0, inf]	-
	mean relative error	MRE	[-inf, inf]	to low values
	mean squared relative error	MSRE	[0, inf]	to low values
	relative volume error	RVE	[-inf, 1]	-
Combination of errors	Kling-Gupta efficiency	KGE	[-inf,1]	-

vide that information to decision makers, requires both information of the climatological anomalies as well as an understanding of the underlying hydrological systems.

The persistence of a system is a measure of how a hydrological condition at a certain point in time can influence the following period and can also be seen as the memory of the system. Catchments with a small storage usually have a small persistence while catchments with large storages can have longer persistences. The predictability of streamflow and other hydrological variables is highly connected to persistence and there exist various methods to estimate persistences. A classical approach to estimate short term persistence is to calculate the autocorrelation of the time series of streamflow observations (e.g. Vogel et al., 1998; Pagano and Garen, 2005). Applying the autocorrelation to highly seasonal data like streamflow data means that they first need to be de-seasonalized before a signal from the autocorrelation can be found other than the seasonality. De-seasonalization procedures for hydrological data, however, often require calibration themselves, as the seasonality rarely corresponds to calendar dates (Hipel and McLeod, 1994).

Several recent studies try to quantify the impact of initial conditions on the predictability of hydrological conditions. Snow cover (Gobena and Gan, 2010; Mahanama et al., 2012), catchment size (Li et al., 2009), North Atlantic Oscillation (NAO), El Niño-Southern Oscillation (ENSO) driven by the Sea Surface Temperature (SST) (e.g. Bierkens and Van Beek, 2009) are generally found to be sources of predictability and they are all highly dependent on the region, system and season. While temperature and precipitation are to parts predictable because of the low-frequency variability in global energy stores, particularly in the ocean (Westra and Sharma, 2010; Feng et al., 2011), on a local scale there are feedbacks because of, for instance, albedo or catchment moisture storages that affect the partitioning between sensible and latent heat flux. Predictability in streamflow is controlled by storages, including snow, soil moisture and groundwater, which attenuate the high-frequency rainfall variability to a lower-frequency streamflow response. Singla et al. (2012) assessed the predictive skill of seasonal hydrological forecast in France with two experiments looking at the influence of land surface initial states on the one hand and atmospheric forcing on the other hand. They focused on the spring season as it is critical to the onset of low flows and droughts. One of their important findings was that the predictability of hydrological variables in France mainly depends on temperature and precipitation in lower elevation areas and mainly on snow cover in high mountains.

1.7 Objective, research questions and approach

The response of streamflow to historical meteorological droughts varied strongly between different catchments. The overarching research question is therefore why is streamflow response so diverse, i.e. what are the main drivers. The main objective is to find these drivers. The research of this thesis approaches the objective from three different angles. Each aspect has its own research questions:

1. **Systematic test of streamflow response to meteorological droughts**

    - How sensitive does streamflow of different catchments react to progressive drying?
    - Can the sensitivity of catchments to drought be detected and quantified?
    - Which catchment properties act as main drivers of drought sensitivity?

    Here, Swiss catchments with different characteristics were systematically tested for their drought sensitivity in model experiments with progressive drying scenarios (Appendix 1).

    - Can a relationship between streamflow persistence of a catchment and its catchment-specific properties be established?
    - Which relative influence do catchment-specific properties have on low flow?
    - Does this relationship vary seasonally?

    The relationship between modeled persistence and catchment properties as well as the relative influence of weather versus catchment-specific properties were investigated in model experiments to elucidate patterns of reactions associated to catchments (Appendix 2).

2. **Consideration of the specific role of snow influence in streamflow response to meteorological droughts**
   Snow influenced catchments show a different discharge regime compared to catchments without snow and also for low flow processes snow can be an important amplifier (Van Loon and Van Lanen, 2012). Particularly in areas that are snow influenced, the recognition of hydrological droughts with common drought indices is unsatisfying (e.g., Van Loon et al., 2014). Therefore, the second aspect is the investigation of snow as potential modifier of streamflow response to meteorological droughts.

   - Does the predictability of droughts change with the degree of snow influence?
   - Do initial snow conditions alter resulting drought predictability?
   - Is a simple index that includes the important snow processes suited for the recognition of hydrological droughts?

   The predictability of droughts is investigated using conceptual models and in particular the initial conditions of snow (Appendix 2). The importance of snow was further investigated by the direct incorporation of snow accumulation and melt in a drought index (Appendix 3). Swiss catchments with different degree of snow influence were investigated to elucidate the suitability of the inclusion of snow processes in the calculation of a drought index.

3. **Representation of different streamflow responses to drought by hydrological modeling**
   There are still several uncertainties related to hydrological modeling that are not yet investigated with focus on low flows (section 1.6). The different responses of streamflow have to be adequately represented and implemented in models. It is necessary to be aware of what has to be considered and what are the particular challenges when modeling low flow.

   - Which model structures are particularly critical or suitable for low flow simulations?
   - How long before a drought occurs could it potentially be predicted?
   - What is the persistence of initial conditions in hydrological models?
   - Does this persistence vary between catchments and time?

   Structures of conceptual models were analyzed to find appropriate model structures for low flow simulations (Appendix 4). The persistence of initial conditions for a common conceptual model was investigated for various catchments to elucidate drought predictability (Appendix 2).

2

Study catchments and data

2.1 Study catchments

For this thesis all in all 29 catchments served as study catchments. Most of the catchments are located in Switzerland and one in Norway. The papers in Appendices 3, 1 and 2 were based on Swiss catchments, whereas the Norwegian catchment was used the paper in Appendix 4. Many of the Swiss catchments that are investigated in the different studies of this thesis are hydrological study areas of the Swiss Federal Office of the Environment (FOEN). Since 1957 the section of hydrology of FOEN maintains several hydrological study areas in order to observe long term changes of the water balance in natural catchments located in different climatic regions of Switzerland. Therefore, detailed descriptions about discharge, areal precipitation, land use, soil, geology and storage capacities as well as several other metrics characteristic for each study catchment are provided. The different climatic regions of Switzerland (Fig. 2.1) represent areas similar climate which is different to the neighboring region's climate. This is of course dependent on the scale used for the comparison (Schüpp and Gensler, 2012). The differences between the climate regions are mainly due to position relative to large scale weather pattern, exposition and elevation of the areas. Since Switzerland is influenced by continental, Atlantic and Mediterranean weather systems that are to different degrees blocked and orographically modified by the Alps and pre-Alps it offers a great climatic variability over an overall rather small area.

The catchments investigated in this thesis are meso-scale (3 to 350 km^2), near natural catchments located in Switzerland (Figure 2.2, Table 2.1 and Table 2.2) and Norway. The mean elevation of the catchments ranges between 480 m a.s.l. and 2400 m a.s.l.. To investigate the main natural underlying processes, only catchments with minor anthropogenic influence were selected, i.e. no catchments with dams, major water extractions or inflow of sewage treatment plants. Additionally, the catchments have, if any, minimal glacier influence and are equipped discharge stations of satisfactory precision during low flow. Due to this constraints many of the existing gauged Swiss catchments could not be included in the thesis, as there is water transferred from one catchment to another, extracted or added from the river. Additionally, in many catchments water is stored

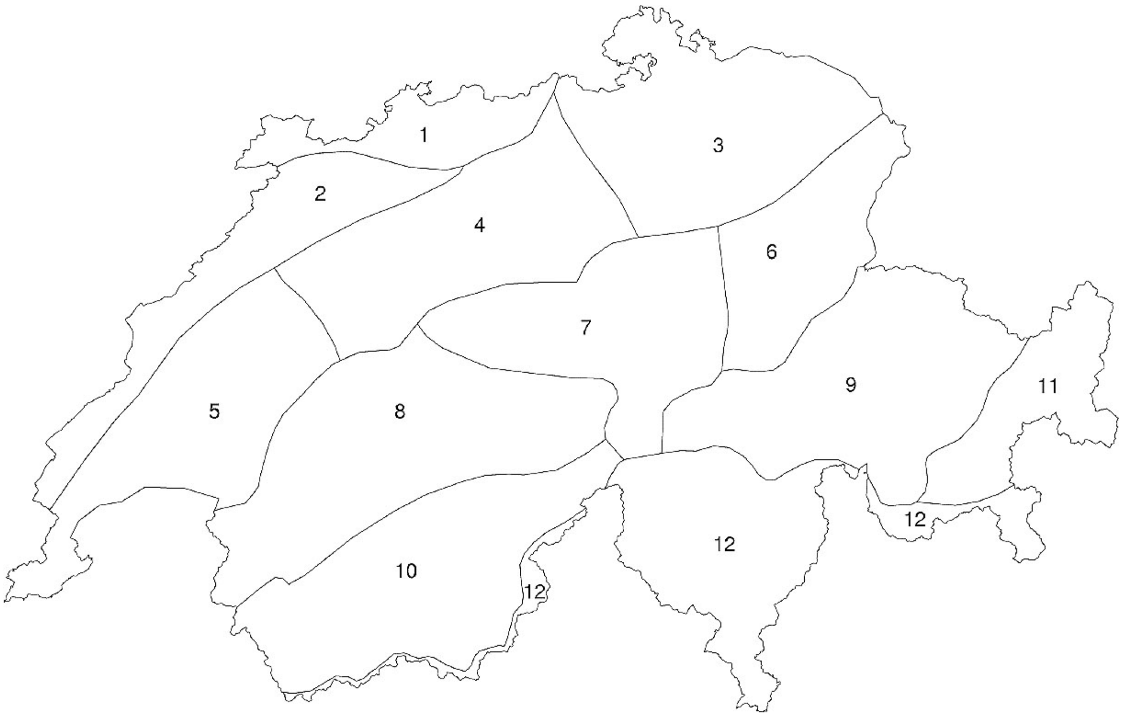

Figure 2.1: Climate zones of Switzerland after Schüpp and Gensler (2012): 1 - eastern Jura; 2 - western Jura; 3 - north-eastern Swiss plateau; 4 - central Swiss plateau; 5 - western Swiss plateau; 6 - eastern part of northern slopes of the Alps; 7 - central part of northern slopes of the Alps; 8 - western part of northern slopes of the Alps; 9 - north and central Grisons ; 10 - Valais; 11 - Engadin; 12 - southern slopes of the Alps

temporally in storage reservoirs. Other catchments simply could not satisfy the measurement precision (section 1.6) that is necessary for low flow studies. The catchments that were selected in the end have all a varying influence of snow (different climate regions see above). Some of the catchments, as of their many elevation zones, have snow processes that occur at different times at different altitudes to a more extreme time lag than others. Table 2.1 and 2.2 give an overview of some general catchment characteristics and of the seasons in which low flow usually occurs.

2.2 Data

The streamflow observations that were used in this thesis were provided by the Norwegian Water Resources and Energy directorate (NVE) for the Norwegian catchment. The meteorological data daily time series of precipitation interpolated from 12 surrounding meteorological stations and potential evaporation was available from Beldring et al. (2003).
The daily streamflow observations that were used in this thesis were provided by FOEN for the Swiss catchments. The meteorological forcing variables for the HBV model, precipitation and temperature, stem from interpolated observations from climate stations (MeteoSwiss) in Switzerland. Before 2013, I used areal hydrometeorological variables

Table 2.1: Catchment information (Engeland, 2002; FOEN, 2012) for the catchments (sorted by increasing mean altitude) that are no study catchments of FOEN the altitude information is based on a digital elevation map of Switzerland (spatial resolution 25m); the calculation of the annual precipitation is based on the period 1975-2012.

Catchment	Size [km^2]	Altitude range [m a.s.l.]	Mean altitude [m a.s.l.]	Hydrological regime	Low flow periods	Annual precip. [mm]	Study used (Appendix)
Aach	48.5	400-600	480	pluvial	Summer	1027	1, 2
Ergolz	261	305-1160	590	pluvial	Summer	1120	1, 2
Aa	55.6	500-800	638	pluvial	Summer	1420	1, 2
Murg	78.9	465-1035	650	pluvial	Summer	1351	1, 2
Mentue	105	445-927	679	pluvial	Summer	1099	3, 1, 2
Broye	392	450-1300	710	pluvial	Summer	1252	2
Langeten	59.9	597-1119	766	pluvial	Summer	1327	1, 2
Rietholzbach	3.3	682-950	795	pluvial	Summer	1341	1, 2
Goldach	49.8	399-1251	833	pluvial	Summer	1429	2
Cassarate	73.9	291-1800	990	pluvial	Summer	1700	1, 2
Guerbe	117	522-2176	873	nivo-pluvial	Winter	1418	1, 2
Biber	31.9	825-1505	1009	nivo-pluvial	Winter	1839	1, 2
Kleine Emme	477	431-2328	1050	nivo-pluvial	Fall, end of winter	1376	1, 2
Ilfis	188	685-2092	1051	nivo-pluvial	Winter, summer	1680	1, 2
Sense	352	548-2189	1068	nivo-pluvial	Winter	1468	3, 1, 2
Alp	46.6	840-1899	1155	nivo-pluvial	Winter	1994	1, 2
Emme	124	745-2174	1189	nival	Fall, end of winter	1702	1, 2
Sitter	261	769-2501	1252	nival	Winter	1913	3, 1, 2
Erlenbach	0.64	1117-1650	1300	pluvio-nival	Winter	2220	1
Luempenen	0.93	1092-1894	1318	pluvio-nival	Winter	2151	1,
Grande Eau	132	414-3185	1560	nival	Winter	1387	1, 2
Simme	344	1096-3217	1640	glacio-nival	Fall, end of winter	1800	2
Schaechen	109	490-3202	1717	regime	Winter	1819	1
Allenbach	28.8	1297-2762	1856	nival	Fall, winter	1659	3, 1, 2
Riale di Calneggia	24	885-2921	1996	nival	Winter	1955	3, 1, 2
Ova dal Fuorn	55.3	1699-3168	2331	nival	Winter	880	2
Ova da Cluozza	26.9	1508-3165	2368	nivo-glaciaire	Winter	964	3, 1, 2
Dischma	43.3	1668-3146	2372	glacio-nival	Winter	1013	3, 1, 2
Narsjø	119	737-1595	945	nival	Winter, summer	594	4

Table 2.2: Geological and hydrogeological characterization of the catchments (sorted by increasing mean altitude) (Engeland, 2002; FOEN, 2012), for the catchments that are no study catchments of FOEN the information comes from the digital vulnerability map of Switzerland (Spreafico et al., 1992).

Catchment	Geology/Hydrogeology
Aach	Molasses, moraines, locally glaciofluvial gravel deposits
Ergolz	Jurassic layers, molasses
Aa	Molasses, moraines, locally glaciofluvial gravel deposits
Murg	Molasses (small permeability); lower part moraines and glaciofluvial gravel deposits
Mentue	Molasses (small permeability) covered with moraines (small permeability) and parts of glaciofluvial gravel deposits (aquifer)
Broye	Sandstones, moraines, flysch, alluvions in valley bottom, locally glaciofluvial gravel deposits
Langeten	Nagelfluh (small to very small permeability), molasses (very small permeability), river alluvions (large permeability)
Rietholzbach	Molasses (small to very small permeability), thin moraine cover, lower part with fissures
Goldach	Molasses (very small permeability), locally moraine (small permeability) with partially higher porosity (storage capacity)
Cassarate	Granite, gneiss, post-glacial till, small alluvions, small part Jurassic
Guerbe	Flysch, molasses covered with moraines and alluvions (very small permeability)
Biber	Molasses covered with thick moraine (small permeability)
Kleine Emme	Small alluvions, sandstones, schists, flysch, moraines, small parts Jurassic
Ilfis	Nagelfluh (middle to small permeability), molasses (very small permeability), alluvions in the valleys (aquifers)
Sense	Small part slightly karstic, molasses, flysch, riverbed locally gravel (very high permeability) else very small permeability
Alp	Molasses (middle to very small permeability); flysch, upper part karstic chalks, alluvions in valley
Emme	Sandstones and schist, moraines, flysch, small alluvions, karstic Jurassic
Sitter	Karstic chalks (very high permeability) to molasses (very low permeability)
Erlenbach	Molasses (middle to very small permeability), flysch
Luempenen	Molasses (middle to very small permeability), flysch
Grande Eau	Flysch, moraines, Jurassic, locally till
Simme	Alluvions in the valley, flysch, Jurassic, schists, locally moraines
Schaechen	Jurassic, moraines, flysch
Allenbach	Mainly 1) flysch with conglomerate and carbonate layers; 2) flysch as well as mesozoic- tertiary carbonates, marl, gypsum and schists; small additional part karstic; high clay content of the dominant rocks causes little to very little permeability
Riale di Calneggia	Granitic gneiss with fissures (variable permeability) talus deposits in the valley (high permeability)
Ova dal Fuorn	Dolomites strongly fissured
Ova da Cluozza	Upper part triadic dolomite and chalk, lower part gneiss
Dischma	Gneiss with fissures, thick moraine layer with high storage capacity, alluvions in the valley
Narsjø	Mainly 1) schists and phyllites in combination with fine grained till soil; 2) igneous rocks combination with coarser till

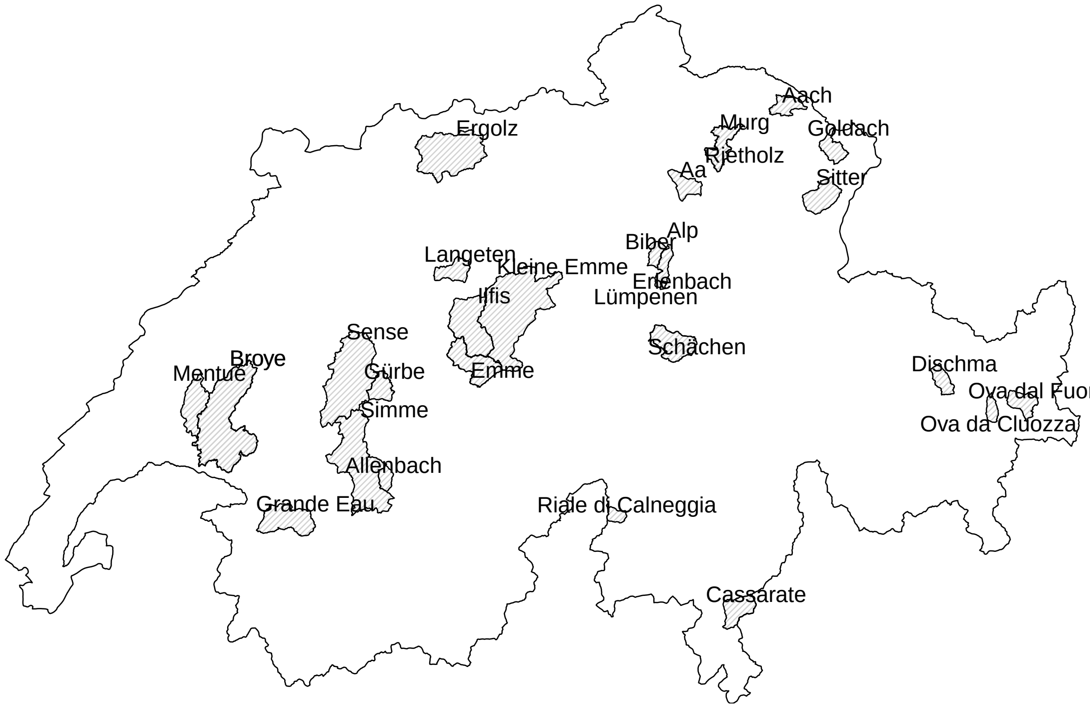

Figure 2.2: Location of the Swiss study catchments.

that were selected from these meteorological stations as well as interpolated and aggregated by the pre-processing tool WINMET (Viviroli et al., 2009), where the spatial and temporal interpolation of observed meteorological variables was based on elevation-dependent regression, inverse distance weighting, Kriging and a simple elevation lapse-rate for temperature data. From this method not only precipitation and temperature but also radiation, wind speed etc. were provided. Hence, with this data potential evaporation could be calculated using the penman formula (Penman, 1948). Longterm monthly values of this potential evaporation served as input for the HBV model.
For studies started after 2013 (Appendices 3 and 1, a grid product of the meteorological variables temperature and precipitation was used, that were provided by FOEN (2013a). Gridded temperature [°C] and precipitation [mm] data (Frei, 2013) available for Switzerland from 1961 - 2013 (MeteoSwiss, 2013) were averaged over each catchment. The interpolation method for precipitation that was used by MeteoSwiss regards topographical effects and is based on a long term averaged climatologic reference field in order to correct for especially the few measurement stations in higher altitudes and thus to minimize the risk for systematic errors (MeteoSwiss, 2013a). Similar to the precipitation also temperature data are homogenized and interpolated with regard to the topography (MeteoSwiss, 2013b).
Table 2.3 shows the periods for each catchment (Chapter 2) that were used in the thesis.

Table 2.3: Data periods of the catchments used in the thesis.

Study	Streamflow	Precipitation	Temperature	Evaporation
A drought index accounting for snow (Appendix 3)	1971-2011	1971-2011	1971-2011	-
Quantifying sensitivity to droughts - an experimental modeling approach (Appendix 1)	1993-2012	1975-2012	1975-2012	1970-2008 (long term monthly)
Comparison of hydrological model structures based on recession and low flow simulations (Appendix 4)	1981-1995	1981-1995	1981-1995	1981-1995 (daily)
Predictability of low flow - an assessment with simulation experiments (Appendix 2)	1970-2008	1970-2008	1970-2008	1970-2008 (long term monthly)

3

Quantifying sensitivity to droughts - an experimental modeling approach*

3.1 Motivation

Meteorological droughts in Europe caused low water levels in lakes, rivers and groundwater. Generally, a prolonged lack of precipitation, storage of precipitation as snow or a strong deficit in the climatic water balance can propagate through the hydrological system causing soil moisture drought and hydrological drought (Section 1.3). The consequences of such droughts are challenging (Section 1.1). Droughts like those in 2003 and 2011 (Section 1.4) are predicted to become more frequent in the future, which calls for a better understanding of the reaction of different systems to droughts. Focusing on single processes in one catchment allows for a detailed analysis of which processes occur during an individual drought event. However, there are not enough observations of historical drought events to perform such a detailed analysis for several events and catchments with resulting detailed links between cause and effect. Historical droughts differed the general preceding wetness and often occurred not at precisely the same time, which makes a spatial and temporal analysis extremely challenging. The diversity of drought propagation mechanisms (Section 1.3) is reflected in the observed droughts, where not every region and catchment was affected similarly in severity and manner. Previous studies looked at historical droughts and tried to link the occurrence and temporal development of a drought with climate and catchment characteristics. Periods of prolonged streamflow drought were found to be caused by the persistent occurrence of specific circulation patterns, however, no clear link between temporal streamflow drought development and observed climatic drought was found.

*Summary of study 1, the full paper (Appendix 1) is submitted to Hydrological Earth System Sciences as: Staudinger, M., Weiler, M., and Seibert, J., Quantifying sensitivity to droughts - an experimental modeling approach.

Many studies have used scenarios to estimate the impact of climate change on streamflow in general and some that focus on droughts in particular. The usual approach is to use simulations of general circulation models or regional climate models with plausible scenarios of greenhouse gas emissions to drive hydrological models. However, large uncertainties are connected to those simulations, and the range of resulting impacts is accordingly high. This study focuses instead on systematic changes. Here, is addressed how sensitive different catchments are to meteorological droughts and whether this sensitivity can be linked to a specific type of catchment, classified by catchment characteristics.

3.2 Methods

HBV modeling experiment

For the modeling experiment the HBV model with the version HBV light was used. Parameter uncertainty was addressed by performing 100 calibration trials, which resulted in 100 optimized parameter sets according to a combination of Nash Sutcliffe model efficiency and volume error. One simulation was run for each of the parameter sets over the entire meteorological observation period and the simulation results of this ensemble of the 100 selected parameter sets were averaged at each time step to derive the reference simulation. The same procedure was performed for the scenarios.

Scenario construction

Two precipitation time series were constructed as hypothetical scenarios with progressively drying conditions:

- Scenario with sorted years (SoYe): All years over the meteorological observation period were sorted from the wettest to the driest year according to the total annual precipitation. Thus, a scenario of modest but continuous progression of drying was constructed.

- Scenario with sorted months (SoMo): For this scenario the individual months were shuffled, with the wettest January together with the wettest February, and so on forming the first year. The second wettest individual calendar months composed the second year. With this approach a scenario was created with a continuous progression of drying in a more extreme manner than SoYe, but still keeping the natural seasonality.

The daily air temperature matching the precipitation from the original time series was re-arranged in parallel to the precipitation scenarios.

Relative change to long-term conditions

First, the relative change of each scenario year to the long-term mean of the reference simulation was calculated for simulated runoff, simulated soil moisture storage, and the combined simulated upper and lower groundwater storages of the HBV model. Second, to assess the catchment sensitivity to the progression of drying the inter-quartile range

(IQR) of the relative change for all variables was calculated. IQR represents the variability during the drying phase and since the scenarios force progressive drying over the course of the years, it can be seen as a measure of sensitivity to droughts: the smaller the value of IQR, the less sensitive a catchment is to droughts, and the higher the value of IQR, the more sensitive a catchment is to droughts. This sensitivity results from both the local climate variability and modification by specific catchment characteristics. Thus, the relative influence of the inter-annual variability of precipitation in each catchment on the scenario was considered. For each year the ratio between mean annual precipitation and long-term mean annual precipitation was calculated and the IQR values were divided by the interquartile range of the precipitation ratios.
The extreme dry end of each scenario was additionally compared to the driest year from the reference simulation in order to determine in which seasons the strongest effect of drying was found. Moreover, the correlations between specific catchment characteristics and sensitivities were analyzed.
Drought characteristics were targeted more specifically, by considering Q_{90} exceedance days per year and calculating a relative change based on the exceedance days. Other indices describing the influence of the progression of drying at its extreme dry end considered in this study, are the ratios of the mean of the driest year of each scenario and the long-term mean for both scenarios. The smaller these indices are, the more sensitive the respective catchments are to droughts.
Further was simulated, how much more each catchment would have been affected if the preceding months to the 2003 drought event would have been drier than in the actual observation. A further index was calculated describing the sensitivity of the catchments to these drier initial conditions. The smaller this index is, the less sensitive the respective catchments are to droughts.

3.3 Main outcomes

Even the modest drying scenario led to a continuous reduction of streamflow, soil moisture and groundwater storage and revealed catchments that were more sensitive to droughts than others. With the more extreme scenario the picture became even clearer. However, for the duration of days above the Q_{90} threshold, an effect was only visible after applying the more extreme scenario. The driest year of the moderate scenario showed seasons with lower than long-term mean streamflow values, that differed for catchments with different streamflow regimes: for the higher catchments where the snow component needed to be considered, there were again higher streamflow values visible in late summer. This could be due to snow melt water filling the storages in spring, which kept the storages at a higher level than would be possible if only rainfed. Further differences between the catchments with nival regimes might be due to different storage release characteristics. This could be confirmed by the analysis of the drought in the summer of 2003 compared to a scenario with drier initial conditions since the storages for the different catchments contributed in different proportions to the reduced streamflow under drier initial conditions.
Comparing the drought sensitivities to catchment characteristics revealed that for both streamflow as well as duration of days above the Q_{90} threshold, mean catchment elevation, size and slope were the main controls. While size was an important predictor in

previous studies elevation improved low flow prediction only in a few studied regions. Soil moisture drought was controlled by size and slope only and for groundwater drought was sensitive to elevation and slope only. Hence, the variability of the storages is not controlled by the same catchment characteristics as the resulting streamflow. However, streamflow comprised all the controls of the storages. The fact that mean catchment elevation is important for drought sensitivity in streamflow can be partly explained by snow in higher elevations. Other reasons like larger storages in higher elevation catchments are indicated by the relationship between groundwater storage and mean catchment elevation.

From previous studies hydrogeology could be expected to be correlated to a storage dependent drought sensitivity, however here no relationship was found. It could be that the hydrogeological productivity number was not an appropriate measure for storage and release or that other controls dominated and hence secondary effects like geology or land use did not show any correlation.

The results that are derived from the modeling experiment contain potential sources of uncertainty, i.e. mainly the choice of the hydrological model and its associated structure and parametrization. The uncertainty from the model parametrization was addressed by an ensemble approach, which generated a more robust simulation than would have been the case for single "best" parametrization. For the model structure can be assumed that the main indication of the results of the streamflow simulation should be similar for different conceptual hydrological models, whereas some differences in the simulated storages can be expected.

The scenarios were constructed by applying sorted annual or monthly precipitation, while air temperature was not explicitly considered. Other studies considered air temperature to climate change by using scenarios with increased temperatures, but constant precipitation. However, the results of previous case studies considering total streamflow response to changes in precipitation and temperature indicated that future total streamflow is more sensitive to precipitation than to temperature. For the construction of the scenarios the preceding wetness of the season was not considered while sorting. This could lead to actual drier or wetter initial conditions for the following year than indicated by the annual sum, particularly for the SoYe scenario. This effect should be minimized by using hydrological years and not calendar years. Still, there could have been a dry summer in an otherwise relatively wet year which then serves as initial conditions for the following year.

The scenarios that were used did not aim to be realistic. Other studies keep the natural variation of precipitation from year to year. Instead the scenarios in this study were constructed to get an idea of how strongly a catchment would react to a moderate and to an extreme progression of drying in comparison with a sample of other catchments from the temperate humid climate of Switzerland. The scenarios were also derived in order to better understand how strongly initial conditions affect hydrological droughts, and were appropriately constructed for this purpose.

As a next step it would be interesting to do an analysis similar to the one in this study for additional regions to find a system of general drivers that make a specific catchment vulnerable to droughts or not. Generally, for water resource management it is important to look at both streamflow sensitivity and storage sensitivity to droughts. With the model-based approach of this study the sensitivity of both can be easily estimated. This approach can serve as a starting point for water resources managers to understand the vulnerability of their catchments. In addition to the used scenarios, scenarios could be constructed with time fractions for sorting that are in between yearly and monthly, for example, scenarios using half a year, a quarter of a year or two months.

This study demonstrates that hypothetical scenarios can be used to evaluate the sensitivity of catchments to droughts. For this, the reaction of streamflow as well as soil moisture and groundwater storages to a continuous progression of drying in general as well as focused on drought characteristics and on one historical drought event were analyzed. The analysis showed that mean catchment elevation, size and slope were the main controls on the sensitivity of the catchments to drought. The results suggest that higher elevation catchments with steeper slopes were less sensitive to droughts than lower elevation catchments with less steep slopes.

4

Predictability of low flow - an assessment with simulation experiments*

4.1 Motivation

The basic objective of drought early recognition is to provide timely warning, so that damages can be reduced or even avoided. The severity of a drought depends clearly on the climatological deficit of water, but also on the hydrological system that has to cope with this deficit. There were many attempts to quantify droughts by indices based on meteorological variables, where each index has its own strengths and weaknesses. Drought indices based on meteorological variables are important, but not sufficient to describe and understand the severity of a hydrological drought. Hence, to recognize locally critical conditions early and provide that information to decision makers, requires both information of the climatological anomalies as well as an understanding of the underlying hydrological systems.

The persistence of a system is a measure of how a hydrological condition at a certain point in time can influence the following period and can also be seen as the memory of the system. Catchments with a small storage usually have a small persistence while catchments with large storages can have longer persistences. The predictability of streamflow and other hydrological variables is highly connected to persistence and there exist various methods to estimate persistences.

Several recent studies try to quantify the impact of initial conditions on the predictability of hydrological conditions. Snow cover, catchment size, North Atlantic Oscillation, El Niño-Southern Oscillation driven by the Sea Surface Temperature were previously found to be predictors and they are all highly dependent on the region, system and season. Predictability in streamflow is controlled by storages, including snow, soil moisture and

*Summary of study 2, the full paper (Appendix 2) is published as: Staudinger, M., and Seibert, J. (2014). Predictability of low flow – An assessment with simulation experiments. Journal of Hydrology, 519, 1383-1393.

groundwater, which attenuate the high-frequency rainfall variability to a lower-frequency streamflow response.
This study looks at the predictability of streamflow with focus on low flows in Switzerland using a conceptual hydrological model (Section 1.6). Thanks to the computational efficiency of conceptual models they can also be used in ensemble prediction systems.
In this study the HBV model was used to perform streamflow simulation experiments and to answer the following questions: How long is the persistence of the initial hydrological state in model simulations of streamflow and does it vary in space and time? Can the persistence be attributed to catchment storage?

4.2 Methods

To quantify the persistence of current hydrological states in streamflow and the influence of weather during prediction three model experiments were designed using the hydrological model HBV in the version HBVlight.

Model calibration

An ensemble of 100 parameter sets was generated for each catchment, based on 100 calibration trials. The mean absolute relative error served as the objective function for the calibration, as the emphasis was on low to medium flows.

Estimation of persistence and catchment relaxation

The first two experiments a) and b) were set up much like the ESP and ESP_{rev} approach. However, in this study 100 parameterizations were used for each ensemble member, which allows a more robust interpretation by using the ensemble mean as well as quantification of parameter uncertainty effects. Experiments a) and b) evaluate both the influence of initial conditions and weather during prediction on the prediction skill.
The simulation experiments differed in the time series that were used as warming up periods to derive initial conditions, and the time series that were used during the prediction period. In experiment a), during the warming up phase the model was forced with different meteorological time series and the forcing during the prediction was the climatology for all simulations. Experiment b) was the reversed version of experiment a); the time series had identical initial conditions, stemming from the climatology. In the simulations, the HBV model was forced with different meteorological time series to derive 'predictions'. For both experiments reference runs were performed: in experiment a) the long term mean was used for both warming up and simulation, in experiment b) the same year as in the experimental run was used for the simulation and the previous chronological year was used for the warming up period. The comparison of simulation and reference runs from experiment a) and b) allowed to estimate streamflow persistences.
A third experiment c) was designed to distinguish further between the influence of the catchments themselves and the meteorological conditions. A relaxation time for the catchments was calculated, defined as the time needed for the system to reach a new

equilibrium after being brought off balance. The warming up was the same as in experiment a). The forcing during the simulation was kept constant and the average annual daily precipitation, mean annual temperature and zero evapotranspiration were used. The precipitation was then distributed to correspond to realistic conditions with precipitation on about 30% of the days. Before running the simulations the initial snow conditions were all set to zero. This was done to remove the influence the melting of accumulated snow had on the relaxation time estimation. Hence, the catchment relaxation time in this study is the streamflow persistence under constant meteorological forcing. The relaxation time [days] was the start of the simulation from experiment c) to the point of an equal oscillation of all simulations.
All experiments a), b) and c) were repeated four times with a shift in the starting date from winter to spring, summer and fall. The starting date is the time where the initial conditions are set, i.e., the switch from warming up to prediction mode. All analyses were performed for each of the 100 parameter sets and for the persistence estimation as well as the catchment relaxation aggregated to a mean value in the end.

Importance of initial conditions vs. weather

The "prediction skill" of both experiment a) and experiment b) forecasts were calculated. As reference, the reference simulation from experiment b) was used as it is the chronologically correct yearly sequence for each forecast/initial condition. Since the focus in the study was on the effects on low flows, the measure of prediction skill of experiment a) and b) was based on the absolute error. The measure of prediction skill of all simulations was calculated for lead times of 1, 2, 3, ..., 52 weeks.

Connection of persistence to conceptual storages

The state of the initial HBV model storages at the start of each simulation were compared to the estimated persistences from experiment a). The actual initial hydrological state at the start of each simulation was transformed to a relative initial hydrological state by using the long term average conditions of the respective month in which the simulation start was set.

4.3 Main outcomes

Hydrological model

From the results of the study the use of an ensemble mean can be recommended, as the variability of the results due to parameter uncertainty was considerable for most of the catchments. The large variability among the simulations that were started in summer and fall when including parameter uncertainty indicates a high uncertainty connected to parameters of the soil routine which control evaporation. The ensemble approach used here is a suitable way to ensure robust results.
Issues such as general model dependency, simulation of snow cover, formulation of potential evaporation have to be considered, also when evaluating the results. However, the main outcomes concerning the influence of initial conditions related to storages within

the catchment are represented and the use of various parameter sets allowed for the estimation of uncertainty derived from the model.

Prediction skill

Previous studies found that depending on when the simulations were started and the lead time applied, the dominance of initial conditions or weather during prediction changed from more dominant initial conditions for short lead times to more dominant weather during prediction for longer lead times. In this study the dominant effect at lead times up to one year was analyzed. At shorter lead times the initial conditions were dominant. Overall the weather as compared to the initial conditions was more dominant for all starting dates. However, the distribution of the ratio of the prediction skill changed for different starting dates. As in previous studies, differences were noted in the ratio of the objective functions for varying dry or wet initial conditions.

Variability of the persistence estimation

The persistence estimates from both experiments overlapped for most catchments. The persistence estimations from experiment b) were systematically longer than the persistences from experiment a) in spring and summer for all catchments. Experiment a) is a representation of what is faced in reality, an attempt to forecast using a known initial condition and several scenarios of how the weather might be. The reference simulations based on the observed weather, allowed to see how long a present/initial state mattered in deriving the most realistic simulation rather than simply initializing the model with the climatology. Instead, in the reference of experiment a) both warming up and forcing used the climatology. The persistences in experiment a) were computationally much faster to estimate than in experiment b) but the climatology plays a greater role in the definition of the persistence estimation. The role of climatology could be the reason for the observed offset in the persistence estimates for the middle elevation catchments: If the initial conditions were wetter during the models warming up phase, it would take longer to reach the reference simulation that was based on a drier climatology than it would take to reach a reference simulation that was based on a realistic seasonal warm up. For fall and winter simulations the climatology was likely closer to a realistic seasonal warm up, since no such offset could be observed.

Streamflow persistence vs. catchment relaxation

The estimated streamflow persistences are a combination of both weather and catchment properties. Catchment relaxation times should instead mainly represent the catchment storage properties. The relaxation times in different seasons, however, can vary slightly as the simulations started with different initial conditions each season and then reached a new balance of the system. The catchment relaxation time for catchments with a snow dominated streamflow regime were longer in spring, which could be explained by filled soil and groundwater storages from the preceding winter and fall. As the lower elevation catchments did not show this seasonal difference the higher catchments are suspected to have larger storages.

Initial conditions and catchment properties

Persistence estimates were strongly correlated to catchment mean elevation. This could partly be explained by an increasing snow influence with elevation, but also by larger aquifers. Initial storages of snow and soil moisture were related to the persistence estimates, a relation that was also found by other studies. Here, the importance of snow, soil moisture and groundwater storage depended on the starting date of the simulations. When the simulations were started in winter or spring the initial conditions of snow were related to the persistence estimates for many catchments and in summer to the highest elevation catchments with more initial snow leading to longer persistences. Drier initial soil moisture could be related to longer persistences for lower elevation catchments with simulation starts in all seasons but winter. For higher elevation catchments and winter simulation start wetter initial conditions led to longer persistences, which might be due to the absolute size of the soil moisture storage of lower elevation catchments compared to higher elevation catchments. The persistences and initial groundwater storage conditions did not show a general pattern.

The role of snow

Snow accumulation is an important storage, snow melt an important storage outflow and it fills other storages in the catchment. This has to be considered when trying to distinguish between meteorological influence and initial conditions with the ESP/ESP_{rev} analysis. Snow melt that contributed to the initial conditions is attributable to the initial conditions, but snow fall, accumulation and melt during the simulation period will directly influence the meteorological forcing. The high elevation catchments, where snow fall could also occur in seasons other than winter, showed a different effect than the catchments at middle elevations, where the initial conditions were still more dominant than the meteorological forcing. This could result from the time shift of when the snow accumulation and melt happened.

For the persistence estimation, snow storage is directly taken into account, which was visible in both the correlation to the mean catchment elevation and the relation between snow storage and persistence. For the catchment relaxation, the direct snow accumulation/melt was explicitly excluded, even though the snow melt that occurred during the warm-up was included. This remaining snow influence seems critical as seasonal differences in the relaxation times of the middle and higher elevation catchments were found, but not in the lower elevation catchments.

Another indication for the role of snow can be seen from the offset between the results from experiment a) and b), namely that the climatology in the warming up of the reference runs in experiment a) were not as realistic as the warm up of experiment b), which caused greater offsets in the seasons with snow involved.

Catchment elevation and storage

At high elevations usually thinner soils are found, however, the results of this study challenge the common assumption of less storage in higher elevation catchments and indicate that there might be a larger groundwater storage, which could be due to large storage features in mountain catchments. The total storage capacity might also increase with elevation because of a storage volume above drainage level that is higher in mountainous catchments than in rather flat low elevation catchments.

Predictability of droughts

The persistence estimations showed that for different catchments the maximum predictability for streamflow varied from 50 days to more than a year with higher elevation catchments related to longer predictabilities. The persistence estimates did not vary greatly with a change of the starting date of the simulations to another season. The relative influence from weather with respect to initial conditions, however, did: in spring the highest elevation catchments had longer lead times with dominant initial conditions presumably due to large snow accumulations at the start of the simulations for all years of the ensemble. At the time of the start of the simulation, the lower elevation catchments have barely accumulated snow, while the snow storage at middle elevation catchments might vary strongly from year to year. This can explain the longer relative influence of the initial conditions on the predictability in the low elevation catchments, but not in the middle elevation catchments, as the snow can accumulate before or after the starting date of the simulation. In fall and winter higher elevation catchments tended to longer lead times of high relative importance of the initial conditions compared to the weather during prediction. This points to a larger influence of the initial conditions in higher elevations which could be due to snow storage as well as other storages. With drier conditions in summer the simulations of catchments at all elevations showed influences of the initial conditions of varying length. The summer ratio of prediction skill point on the one hand to storage differences, but also to varying summer meteorology for the different catchments.

With this study, the question of drought predictability, cannot readily be answered. However, for the catchments in this study ranges of maximum detectable influence of initial conditions were quantified. Further was found that the catchment elevation matters more than the starting date of the simulation for a maximum predictability of streamflow and that the relative importance of initial conditions compared to the relative influence of the weather during the predictions changes with the season in which the simulation start is set.

Further, in opposition to an intuitive expectation from shallow soils in higher elevations, the results indicated for larger groundwater storages in higher elevation catchments. This may motivate a reconsideration of the sensitivity of mountainous catchments to low flows in a changing climate.

5

A drought index accounting for snow*

5.1 Motivation

In contrast to drought processes that occur in summer, droughts caused by the storage of precipitation as ice and snow can act as a key moderator of hydrological drought (Section 1.3). In particular, streamflow droughts are often related to the presence or absence of snow in the preceding winter period and winter droughts can occur despite large amounts of precipitation. Monitoring and managing streamflow droughts requires indicators that are general enough to be easily applicable, but specific enough to capture the type of drought relevant to the region and variable of interest.

The Standardized Precipitation Index (SPI) as well known indicator for drought has been applied in many studies, in operational drought monitoring in the present, and also in scenario predictions of drought for climate change impact assessment. A major advantage of the SPI compared to other drought indices is that it requires only precipitation data to describe drought severity. Different precipitation aggregation periods can reflect the impact of drought as it propagates through the hydrological cycle (Section 1.3).

To create a methodologically consistent indicator of hydrological droughts, several studies have transferred the SPI approach to observed and modeled hydrological variables. The SPI can be modified to move from a precipitation drought characterization to a hydrological drought characterization without the full complexity of a hydrological model by accounting for a known first-order control on catchment hydrology that affects drought. This study aims to find a climatic drought index with low data requirements which nevertheless can serve as an indicator for hydrological drought in regions with a variable influence of snow on catchment hydrology.

*Summary of study 3, the full paper (Appendix 3) is published as: Staudinger, M., Stahl, K., and Seibert, J. (2014), A drought index accounting for snow, Water Resources Research, 50, 7861Ű7872, doi:10.1002/2013WR015143.

5.2 Methods

This study uses a drought index based on observed streamflow against which to compare the climatic drought index SPI and the new Standardized Snow Melt and Rain Index (SMRI). The benefit of the SMRI is demonstrated for seven Swiss catchments with different amounts of snow melt contribution to streamflow.
The new SMRI was calculated similarly to the SPI, but from the daily sum of snowmelt and rain (MR). To obtain daily snowmelt amounts, a simple degree-day method that only requires temperature data in addition to precipitation was first applied. The variable MR was then transformed into the SMRI using the Pearson type III distribution. In order to explore the level of local parameterization needed, three parameter set ensembles were tested: the first parameter set ensemble assumed no prior knowledge, the second set assumed some regional knowledge and the third set assumed specific catchment knowledge. The SMRI was always calculated for different aggregation periods.
The effect of different values for the snow model parameters on the SMRI was analyzed with a Monte Carlo analysis. In addition to a general evaluation of the SMRI, it was also evaluated for two drought events (summer 2003 and spring 2011). The sensitivity of the SMRI to elevation distributions was analyzed by repeating the computations in a semi-distributed way, using an area-weighted mean value of catchment snow melt water from different elevation zones.

5.3 Main outcomes

Uncertainties from model and index standardization

The SMRI is calculated in a two-step process: (1) a model is applied that accounts for the dominant process that affects severity and timing of hydrological drought, and (2) the output of this model is transformed into the index. This two-step process means that the resulting index has multiple sources of uncertainty. The most important sources of uncertainty from the snow model are the parameterization of the degree-day model, the spatial discretization of elevation as well as data uncertainty. Model parameterization and spatial discretization were addressed by ensemble approaches using parameterizations stemming from no prior knowledge, regional knowledge and specific catchment knowledge. The calculation of the actual index is then influenced by semi-objective decisions including that for a theoretical distribution function and finally the choice for an aggregation period.
For the snow influenced catchments and for most parameter combinations, the entire parameter range resulted in an SMRI that was much closer to a hydrological drought description than the SPI for both the entire observation period as well as for the dry periods only. Increasing the knowledge about the snow model parameters of a catchment decreased the uncertainty. However, there was not an increase but a slight decrease of the performance found, which might be explained by the fact that the prior knowledge parameters were derived by calibration of a full hydrological model. The optimal snow parameter values derived in this way might be model specific and not be those providing best results for the SMRI, when soil or groundwater were not considered. These results indicate that the use of an ensemble of random parameters actually might be the most

appropriate approach after all. Overall, the parameterization of the snow model has only a minor influence on the systematic performance of the SMRI.
Using elevation zones in the melt model instead of a lumped mean elevation, improved the performance of the SMRI. The resulting reduction of the range of SMRI values can be attributed to the explicit consideration of higher elevations. Still, when no information on the elevation distribution is available the simpler approach of using the catchment mean elevation resulted in the SMRI describing the streamflow conditions closer than the SPI.
The choice of the climate data input will affect the results. The snow model was driven with uncorrected precipitation data, hence there can be errors due to precipitation undercatch - especially in winter when precipitation falls as snow. However, the resulting bias affects the calculation of both indices, the SMRI and the SPI. Hence, the comparison of the indices and the results presented in this study should not be affected.
The choice to include just one key process, i.e., snow accumulation and melt, in the model used to compute the SMRI has implications for the seasonal performance of the new index. The measure of comparison decreases in the months of May to August for the strongly snow influenced catchments. These are the months with the highest evapotranspiration, a process that was not modeled here, but could be considered. For many snow-dominated catchments the performance gain by including processes other than snow is expected to be small. Despite the exclusion of evapotranspiration, over the entire year the SMRI was closer to streamflow conditions than the SPI and particularly so in the months of snow melt. For the catchments with a pluvial regime, the difference between SPI and SMRI as an indicator for streamflow drought conditions is small or not existent.

Application potential

Similar to other existing drought indices, the new index can be calculated for different aggregation periods. The evaluation during two recent drought events showed that with an increasing aggregation period, the SMRI and SPI approximate each other for the studied catchments. The SMRI is thus considered useful to indicate streamflow droughts, that occur in humid to semi-arid mountain regions on a time scale below one year due to the seasonal character of snow storage.
In the temperate humid climate of Switzerland, snow melt and precipitation occur in the same season, as rainfall is uniformly distributed over the year. This requires shorter aggregation periods to be considered for the calculation of the indices than in a Mediterranean climate, where wet and dry seasons are clearly separated. Where such a clear separation does not exist, other indices that include end-of season snow pack directly will be less useful for drought assessment.
Ideally, an index also needs to be suitable for regional comparisons, i.e., easily applicable with distributed or gridded climate datasets and without further information needs. Validation of the SMRI approach with a standardized streamflow index from streamflow observations in meso-scale catchments across a gradient of increasing snow influence, as proposed in this study, shows that for these cases the simpler approach is a suitable alternative to more complicated or limited approaches to describe the evolution of hydrological drought situations.

The SMRI, as introduced in this study, combines the low data requirements of the SPI with the explicit consideration of snow accumulation and melt. The analyses of the new index demonstrates its usefulness to indicate hydrological droughts in snow influenced catchments, with specific advantages in those climatic regions where snow melt and rainy season coincide. This case study with Swiss catchments suggests a closer description of hydrological droughts by the SMRI than by SPI. Following the gradient of snow influence, the more a catchment is influenced by snow the more worthwhile it is to complement the SPI with the SMRI.

6

Comparison of hydrological model structures based on recession and low flow simulations*

6.1 Motivation

Low flow and droughts affect many sectors (Section 1.1) and occur in every country albeit in different perceived severity. Thus monitoring and modeling of low flow are crucial for their analysis and prediction. However, low flows are poorly reproduced by many hydrological models since these are traditionally designed to simulate the runoff response to rainfall.

A revision of model concepts regarding low flows requires a clear understanding of the model's structural deficits. A common approach to investigate the impact of the differences in model structure is to perform model inter-comparison experiments. Such experiments have been helpful to explore model simulation performance of lumped, semi-distributed and distributed models in a consistent way using the same input data. The reasons for the differences remain unclear since each model uses different interacting parameterizations to simulate the hydrological processes. Discrepancies between observed and simulated streamflow can arise from errors in the input data rather than weaknesses in model structure, which complicates the investigation of the impact of the differences in model structure. The Framework for Understanding Structural Errors (FUSE) is a computational framework that enables a separate evaluation of each model component by modularizing individual flux equations instead of linking available submodels. FUSE identifies the set of subjective decisions while creating a hydrological model and offers multiple options for each model decision. This approach can thus help to get a better understanding of the occurring hydrological processes.

Commonly, streamflow recession is modeled as the outflow from a, or a set of, linear or

*Summary of study 4, the full paper (Appendix 4) is published as: Staudinger, M., Stahl, K., Seibert, J., Clark, M. P., and Tallaksen, L. M. (2011). Comparison of hydrological model structures based on recession and low flow simulations. Hydrology and Earth System Sciences, 15(11), 3447-3459.

non-linear reservoirs (Section 1.5.3). In periods with no input outflow from the reservoirs control the streamflow and thus, the model behavior during low flow. Real hydrological processes can be more complex. Thus, it is of interest to have a closer look at the hydrograph recession, and carefully evaluate model simulations of recession behavior.
Several studies exist that link recession analysis with the structure of hydrological models (Section 1.5.3). In this study the model structures are systematically analyzed using FUSE. The associated model performance is evaluated with respect to the ability to simulate low flows and recession behavior. This is done for one catchment only to allow a more detailed insight into the model structures. The main objective is to investigate the relative influence of a single model structure on the model performance. As there are distinct differences in the recession rates found for summer and winter, one task is to study how model structure is connected to the seasonal performance for low flow simulation.

6.2 Methods

Snow accumulation and melt

Narsjø is a snow dominated catchment (Chapter 2), however, there was no snow routine implemented in the version of FUSE used for this study. Hence, the input data was pre-processed with a snow accumulation and melt model which corresponds to a single implemented snow routine. Here, a simple degree day method was applied.

FUSE framework

The use of FUSE as a diagnostic tool to detect the impact of model structure involved the following three steps: (1) prescription of the type of model (2) definition of the major model-building decisions and (3) preparation of multiple options for each model building decision. In this study, the type of model was limited to lumped hydrological, that were run at a daily time step. Four conceptual parent models were selected to be recombined to new FUSE-models. The selection of the parent FUSE models was limited to four well known models, covering common principles used in conceptual hydrological models. All parent models consist of equally plausible structures and the important processes could be broken down into fluxes occurring in the upper layer and lower layer, evaporation, percolation, subsurface flow and surface runoff (model building options).
Some processes were not explicitly modeled, including interception by the vegetation canopy as well as specific surface energy balance calculations. Routing was calculated by a two parameter Gamma distribution. Thus, all models represent the subsurface with a similar level of detail and thus differences that emerged from different plausible model structures were emphasized rather than differences due to the set of processes represented.
The model decision options that were made separately for each of the FUSE models with regard to the upper soil layer, lower soil layer, evaporation, percolation, subsurface flow, surface runoff, bucket overflow and runoff routing.

Model calibration

All FUSE models were calibrated using the Shuffled Complex Evolution algorithm. The mean absolute relative error was chosen as objective function, as it emphasizes low to medium flows.

Low flow and recession analysis

The performance of the model was evaluated using the logarithmic Nash-Sutcliffe efficiency, which also emphasizes model performance in the low flow range. In this study the relationship between the negative change in streamflow over time $-\frac{dQ}{dt}$ [mm day^{-2}] and the corresponding streamflow Q [mm] was analysed. For the evaluation of the model performance of recessions both modelled and observed data were used. As both $-\frac{dQ}{dt}$ and Q span several orders of magnitude, their relation is plotted in log-log-space. All plotted points in a certain range of Q were averaged to one value representative for this range. Then, a polynomial function was fitted to the relationship between $-\frac{dQ}{dt}$ and Q. The polynomial coefficients were fitted using a least squares regression model. Scatter plots of the polynomial coefficients were then used to compare observed and simulated recession behavior for the FUSE models. The recession behavior was analyzed for both the whole year and summer and winter seasons.

6.3 Main outcomes

The basic assumption in this study was that different model structures are the reason for the differences in model performance. Only four models performed well regarding the logarithmic Nash Sutcliffe efficiency for both the whole year and for summer and winter. All used a combination of certain lower layer/subsurface flow, upper layer and the percolation, containing at least two of the three components. For all other well performing models a systematic influence of a specific structural decision could not clearly be found. Structural decisions that cause poor performance could be tracked based on the performance criteria and the simulation of the recession relationships. Such a structural decision is a certain combination of lower layer/subsurface flow with percolation. This combination caused poor low flow simulations for the whole year as well as for the seasonal time series. The comparison of the slopes of summer and winter recessions reveals no seasonal differences for models with exactly this lower layer/subsurface flow and percolation combination. The lower layer corresponds to the subsurface flow which is conceptualized by a power law originating from the parent model TOPMODEL. The main difference between the subsurface flow parametrization in TOPMODEL and the other parent models is its dependency on the underlying distribution of the topographic index. The storage capacity in TOPMODEL also depends on the topographic index distribution and can hence be smaller or larger depending on the topography. In this study the Gamma distribution was used , which is generally considered to be an appropriate assumption for the topographic index distribution of most catchments. However, the models that used the TOPMODEL options may not have represented the topography in the Narsjø catchment well enough.
The specific percolation option is dependent on the lower layer decision. It thus strength-

ens the assumptions made with the lower layer/subsurface flow decision. This percolation option causes the fastest drainage when the lower layer is dry. Steep recession slopes were modelled with this specific combination of percolation and lower layer/subsurface flow. The calibration with this combination appears to have caused a small water holding capacity of the lower layer resulting recessions that are steeper than recessions in the observed data.

Model input in winter is precipitation plus snow melt. Towards the end of the winter season (May/June) small snow accumulations and melt events might fill the storages with small amounts of melt water and produce a prolongation of the recessions. The recessions modeled with the specific combination of lower layer/subsurface flow and percolation options are too fast and this results in unrealistic shapes of the recession relationships. The specific percolation option hence seems inappropriate for a combination with the specific lower layer/subsurface flow as it results in recessions that are too fast in summer and in streamflow that are too low in winter. None of the model decision combinations has such a distinct influence on model performance as this specific combination.

There are further combinations that systematically influence the seasonal performance: according to the results the assumption of a percolation based on the field capacity should not be used to simulate winter recessions. In summer, however, other model decisions cause a poor performance than in winter: one example is the surface runoff structure from TOPMODEL that models poor summer recessions. In summer, surface runoff plays a larger role for recessions than in winter.

Generally, model performance for low flows is easier to analyse for winter than for summer. In summer, there are several fast responding storages that contribute to the streamflow. The longer the recessions last, the less important become quickly draining storages that are prone to evaporation while slowly draining storages gain more influence. In addition, there can be a considerable influence by transpiration In winter, the only storages that are important are lower layer storages and snow. Since only one snow storage option was modeled, only the lower layer storages matter. The results point out that the most important features for winter recession are directly connected to the lower storages. Hence, it is rather surprising to find a distinct modeling decision that causes a similar performance for both winter and summer recessions.

In this study the choice of model structures was constrained to the structures of only four parent models. To keep the analysis manageable, some processes were explicitly exempt. This includes climate input and hence required the preprocessing of the input data was with a snow accumulation and melt model instead of including several structural decisions of a snow model. Snow is in fact important in the Narsjø catchment, and testing structures describing the processes related to snow might be worthwhile. This study, however, focused on the impact of model structures used to represent groundwater storage and release behavior. Future applications should consider testing more structures describing processes of snow melt and accumulation, but also interception and evapotranspiration, all of which were described with a single structural decision in this study. Further, the storage structural decisions included in this study are not the only options. Combinations of linear and non-linear reservoirs in series or parallel could be appropriate for the Narsjø catchment. In general, validation of models with observed data of poor quality may lead to the rejection of models that might in fact be appropriate. A way to avoid the evaluation of model performance by standard metrics, such as the mean squared error, diagnostic signatures can be used. To include additional data on individual processes within a catchment may be necessary to identify scientifically defensible modeling strategies.

This study assessed the impact of model structure on low flow simulations and recession behavior using FUSE. Specific model structure combinations of different conceptual models resulted in different model performance for summer and winter low flows. Overall, individual structural decisions never appeared to be an exclusive reason, but rather the combinations of specific structural decisions affected model performance. Evaluating with the logarithmic Nash Sutcliffe efficiency as objective function, led to only a small number of models that performed well. While most well performing models did not allow for the detection of a systematic influence of a model structure combination on the model performance, poor performance was more clearly linked to specific model structures.

An important task would be to test this further for additional catchments with a seasonal flow regime. In order to elucidate to which extent the influence of the considered model on low flow simulations are catchment specific or can be generalized, it should be replicated in other catchments.

7 Synthesis

The overarching question of this PhD thesis was why streamflows of different catchments respond differently to meteorological droughts. The research was dedicated to give answers to this question approaching from three different angles.

7.1 Systematic test of streamfow response to meteorological droughts

In the thesis a model experiment was designed using scenarios with progressive drying (Appendix 1). That way the sensitivity of catchments to meteorological droughts could be evaluated. For the studied catchments mean catchment elevation, size and slope were found as main drivers for the sensitivity of streamflow to meteorological droughts. The results suggested higher elevation catchments with steeper slopes were less sensitive to droughts than lower elevation catchments with less steep slopes. High elevation catchments also revealed longer persistence estimates (Appendix 2) than lower elevation catchments, hence both persistence estimates and drought sensitivities point at larger storage connected to higher elevation. From these results could be concluded that one or more drivers are related to mean catchment elevation, for instance, snow accumulation (section 7.2). However, the correlation of sensitivity and persistences with mean catchment elevation could also stem from other drivers that only in the selection of the catchments incidentally indicate some relationship to elevation. This was tested by filtering the snow influence and analyzing a catchment relaxation (Appendix 2). To elucidate this relationship further it would be worth testing additional catchments and look at the link between other storages than snow and catchment elevation.

While the studied catchments could be ranked in their drought sensitivity within the group of catchments, the quantification of drought sensitivity can still be improved. The method allows to break the sensitivity down into numbers and, as in this thesis, regional comparison is possible. The more catchments would be included in such an analysis, the more meaningful should the quantification be, i.e. the sensitivity would be sort of calibrated.

To find drivers, it often is worthwhile to be able to distinguish between the influence of weather and of catchment-specific properties like groundwater storage, soil properties etc.

for analysis. However, both are tightly connected (e.g., Peters et al., 2003; Tallaksen and van Lanen, 2004; Van Loon, 2013; Van Loon et al., 2014). Van Loon et al. (2014) looked separately at the influence of climate and groundwater storage by using synthetic catchments in various climatic regions. While this is an elegant approach to see a qualitative behavior, to quantify the contribution of catchment-specific properties of real catchments this approach might not be sufficient. The model experiments (Appendix 2) helped to understand the relative importance of initial conditions. Drier initial conditions of soil moisture and groundwater as well as more initial snow resulted in longer influences of initial conditions.

A complicating fact for the interpretation of the results is, that the initial conditions originate only in part from catchment properties. Most storage features can said to be catchment-specific, but their filling not exclusively. Instead of looking qualitatively it would be worth to systematically investigate the relation between certain catchment-specific properties and initial conditions and quantify their influence.

It is even more difficult to distinguish between the contribution of weather and catchment properties to hydrological drought development as snow acts as a water storage when it is accumulating and as a water source when it is melting (Appendix 2). In the thesis the effect of snow was minimized by computing a relaxation time of the catchment.

7.2 Consideration of the specific role of snow influence in streamflow response to meteorological droughts

All catchments that were involved and investigated in this thesis are to a certain degree snow influenced. The results showed, how important snow processes are in snow influenced regions and that they should be considered for both hydrological drought indices (Appendix 3) and low flow modeling (Appendix 1, Appendix 2). Snow processes, as key drivers in snow influenced catchments, should not be neglected in drought indices as was stated by Van Loon et al. (2014) and was shown here by the $SMRI$ that specifically includes snow accumulation and melt (Appendix 3). The results of the thesis indicated the importance of snow for predictability of low flow and sensitivity of catchments to meteorological droughts (Appendix 1 and Appendix 2). Longer streamflow persistence could be found in higher elevation catchments (Appendix 2), dependent on the initial snow conditions. Catchments with higher snow influence appeared less sensitive on meteorological droughts, even after accounting for precipitation differences between the catchments (Appendix 1).

A key result from analyzing model structures for low flow modeling (Appendix 4) was that some model structures performed better when evaluating only summer or winter, which again points at the importance of snow. In winter other processes contribute to low flow than in summer (e.g. Burn et al., 2008) which has to be included in the model structures.

7.3 Representation of different streamflow responses by hydrological modeling

Structural errors should be considered in hydrological modeling as they can be large but very critical with regard to potential future applications, particularly for prediction (e.g., Beven, 2001; Solomatine and Wagener, 2011). In the thesis model structures of different

conceptual hydrological models were analyzed to their suitability for modeling low flow (Appendix 4). Some structures were found to be not suited for the studied catchment. Further, a seasonal difference in the model performances was found. The importance of structural decisions (Clark et al., 2008) was confirmed also for low flow, where particularly groundwater structures were found to be critical.
An important task would be to test model structures for low flow modeling further for additional catchments with a seasonal flow regime. Further diagnostics other than used in this thesis (e.g., Wagener et al., 2007; Sawicz et al., 2011), could be used to evaluate the performance of structures for low flow modeling to get a more comprehensive view. Ideally would be to include diagnostics that are not based on streamflow observations, for example, the comparison with changes in the isotopic composition in streamflow (e.g., McGlynn et al., 2003; McGuire et al., 2005; Tetzlaff et al., 2009).
The maximum predictability could be estimated with model experiments (Appendix 2). The results indicated that the persistence of initial hydrological states was varying with mean catchment elevation, i.e. for higher elevation catchments longer persistence were estimated. This could be explained by the role of snow as well as groundwater storages, which were connected to the higher elevation catchments of the catchment selection of this thesis (section 2).
The model experiments helped further to understand the importance of timing and influence of initial conditions. Qualitative statements could be made about the importance of start of the simulation and influence of initial conditions on the predictability. The quantification of a predictability of droughts, however, was not readily solved. For a better quantification, the method could be combined with some diagnostic signatures (e.g., Sawicz et al., 2011) specifically targeting low flow.
It was analyzed how long various initial hydrological states are persistent in model simulations of streamflow and if the resulting persistence changes significantly with a simulation start in different seasons (Appendix 2). The persistence estimates did not vary greatly with a change of the start of the simulations to another season. The relative influence of the initial conditions with respect to weather, however, varied considerably with a change of the simulation start.
While the performance of modeling low flow was compared by looking at individual model structures from common conceptual hydrological models (Appendix 4), the predictability of droughts and the sensitivity of different catchments to droughts was tested for the HBV model only (Appendix 2 and 1). Potential differences occurring from the use of different models could be investigated in a future study.

7.4 Groundwater storage

Less clear than snow, also groundwater storage was found to be an important driver: both the results of the thesis indicate a longer predictability and less sensitivity for higher elevation catchments (Appendix 1 and Appendix 2). This could, however, not solely be attributed to more snow. Tague and Grant (2009) suggest a key role of subsurface properties for streamflow sensitivities next to snow and modifying the role of snow. A conclusion from the thesis (Appendix 1 and Appendix 2) was to consider larger storage capacities for higher elevation catchments due to large aquifers. The results (Appendix 2) indicated that larger groundwater storage capacities lead to longer persistence

and thus predictability of low flow. Appendix 4 showed that poorly chosen groundwater structures can lead to poor model performance with regard to low flow, which calls for a reflection of the model structure before applying models for purposes such as drought early warning.
Even though the results of the thesis indicated an influence of groundwater storage on the sensitivity of catchments to meteorological droughts (Appendix 1), there was no correlation to the used measure of hydrogeology. Similar measures for hydrogeology derived from GIS computations were tested by Kroll et al. (2004), but also these measures where not found to be satisfactory. As subsurface properties clearly influence streamflow regimes these properties are tried to be estimated either through calibration of hydrological models or through inferences based on catchment similarity (e.g., Beven, 2006; Wagener and Wheater, 2006; Tague and Grant, 2009). In this thesis as well as in previous studies (e.g., Kroll et al., 2004; Tague and Grant, 2009; Stoelzle et al., 2014a; Stoelzle et al., 2014b) researchers are confronted with the conceptual idea of the control of subsurface soil and geologic properties controlling the rate of water flux on the one hand and a complex geological reality on the other hand. Nevertheless, it would be useful to find in a future study a simple proxy for groundwater storage that can be linked to the sensitivity to meteorological droughts. Then it could be also analyzed, when and where this sensitivity is strongly connected to groundwater storage or when and where weather - and snow in particular - overprints catchment characteristics.
Recession analysis is useful as a diagnostic tool for model evaluation (McMillan et al., 2011; Clark et al., 2011). In this thesis recession was as a signature to evaluate model structure (Appendix 4). The missing link between hydrological drought sensitivity and groundwater by the measure that was applied (Appendix 1) could maybe be closed using a signature derived from recession analysis: Catchments with a slow recession rate are typically groundwater dominated, while impermeable catchments with little storage show faster recession rates (Tallaksen and van Lanen, 2004). Hence, it would be interesting to investigate if a signature from recession analysis could serve as a proxy for groundwater storage at least in terms of reaction time. To be considered in such an approach should be the decision of which part of the recession are included in the evaluation and the choice of method for recession analysis (Stoelzle et al., 2013).
Generally, the role of groundwater for low flow modeling and drought early recognition is still a challenge that is worth investigating. Apart of the recession analysis as potential method to be applied also using tracers such as stable isotopes should also here be helpful to estimate storage capacities and thus drought sensitivities.

7.5 Applications

Drought early warning is an important issue as it can help to warn in an appropriate time and thus improves the management. The here developed $SMRI$, a drought index that accounts for snow, can be a useful tool for drought monitoring and is hence a contribution to an improved drought early warning particularly in snow influenced areas. Since the $SMRI$ is less complicated than a full hydrological model it could be easily included in existing drought observing platforms. As Swiss catchments were used to develop the $SMRI$ it could be directly implemented in the existing drought platform `drought.ch` (Seneviratne et al., 2013).

If more processes should be included, modeling low flow might be required. It was shown that model structures in particular regarding groundwater conceptualization are crucial for the model performance of conceptual models to simulate low flow. Hence, for the application of models, for instance, for drought early recognition, it is advisable to test model structures before application.
Drier initial states of soil moisture and groundwater and more snow accumulation at the start of the simulation led to longer persistence estimates (Appendix 2). For model predictability it is thus important to simulate these as well as possible - which includes appropriate structures - in order to estimate a realistic predictability. Regarding the influence of weather and initial conditions on a “prediction skill”, the contribution of either part changed for different starting dates and for some catchments even more strongly. Wet initial conditions in spring showed a dominant contribution of the weather. With drier initial conditions in summer, the contribution of initial conditions is larger for longer compared to spring. For application this knowledge about effects of rather wet or dry initial conditions can thus help to evaluate the drought predictability.

APPENDIX 1

Quantifying sensitivity to droughts - an experimental modeling approach*

Abstract

Meteorological droughts like those in summer 2003 or spring 2011 in Europe are expected to become more frequent in the future. Although the spatial extent of these drought events was large, not all regions were affected in the same way. Many catchments reacted strongly to the meteorological droughts showing low levels of streamflow and groundwater, while others hardly reacted. The extent of the hydrological drought for specific catchments was also different between these two historical events due to different initial conditions and drought propagation processes. This leads to the important question of how to detect and quantify the sensitivity of a catchment to meteorological droughts. To assess this question we designed hydrological model experiments using a conceptual rainfall-runoff model. Two drought scenarios were constructed by selecting precipitation and temperature observations based on certain criteria: one scenario was a modest but constant progression of drying based on sorting the years of observations according to annual precipitation amounts. The other scenario was a more extreme progression of drying based on selecting months from different years, forming a year with the wettest months through to a year with the driest months. Both scenarios retained the typical intra-annual seasonality for the region. The sensitivity of 24 Swiss catchments to these scenarios was evaluated by analyzing the simulated discharge time series and modeled storages. Mean catchment elevation, slope and size were found to be the main controls on the sensitivity of catchment discharge to precipitation. Generally, catchments at higher elevation and with steeper slopes seemed to be less sensitive to meteorological droughts than catchments at lower elevations with less steep slopes.

*Submitted to Hydrological and Earth System Sciences as: Staudinger, M., Weiler, M., and Seibert, J., Quantifying sensitivity to droughts - an experimental modeling approach. This version contains minor corrections raised by the review to date.

1 Introduction

Meteorological droughts such as the summer drought of 2003 (Rebetez et al., 2006) or the spring drought of 2011 (Kohn et al., 2014) in Europe caused low water levels in lakes, rivers and groundwater. Generally, a prolonged lack of precipitation (meteorological drought), storage of precipitation as snow or a strong deficit in the climatic water balance can propagate through the hydrological system causing soil moisture drought and hydrological drought (Tallaksen and van Lanen, 2004; Mishra and Singh, 2010). The consequences of such droughts are challenging: water use restrictions have to be applied to, for instance, energy production or irrigation. Water quality can be affected by faster warming of less than usual water and reduced dilution, which in turn becomes an issue for ecology, but also for drinking water supply. Droughts like those in 2003 and 2011 are predicted to become more frequent in the future (Solomon, 2007), which calls for a better understanding of the reaction of different systems to droughts. Focusing on single processes in one catchment allows for a detailed analysis of processes occurring or not occurring during a individual drought event (Santos et al., 2007; Trigo et al., 2010; Li et al., 2010). However, there are not enough observations of historical drought events to perform such a detailed analysis for several events and catchments with resulting detailed links between cause and effect. Historical droughts usually differ in initial conditions regarding the general preceding wetness and often occur not simultaneously, which makes a spatial and temporal analysis extremely challenging. A meteorological drought can develop into a hydrological drought through different mechanisms that are controlled by catchment characteristics as well as climate (Eltahir and Yeh, 1999; Peters et al., 2003; Tallaksen and van Lanen, 2004; Van Loon and Van Lanen, 2011): several consecutive meteorological droughts can turn into a combined and prolonged hydrological drought and they can be attenuated by the storages of a catchment. Further there is often a varying time lag between meteorological, soil moisture and hydrological drought that involves both streamflow and groundwater (Van Loon and Van Lanen, 2011). In addition to a deficit in precipitation, droughts can also be caused by temporary storage of water as ice or snow (Van Loon et al., 2010). This diversity is reflected in the observed droughts, where not every region and catchment was affected similarly in severity and manner. Based on the different drought generating mechanisms, Van Loon et al. (2010) developed a general hydrological drought typology and distinguished between six different drought types that include the type of precipitation and air temperature conditions preceding the drought (classical rainfall deficit drought, rain-to-snow-season drought, wet-to-dry-season drought, cold-snow-season drought, warm-snow-season drought, and composite drought).

Previous studies looked at historical droughts and tried to link the occurrence and temporal development of a drought with climate and catchment characteristics such as, for instance, topography or geology (e.g. Stahl and Demuth, 1999; Zaidman et al., 2002; Fleig et al., 2006). Stahl and Demuth (1999) found that spatial and temporal variability of streamflow drought was influenced by the geographical and topographical location and the underlying geology. Periods of prolonged streamflow drought were found to be caused by the persistent occurrence of specific circulation patterns, however no clear link between temporal streamflow drought development and observed climatic drought was found.

Table 1.1: Catchment characteristics (FOEN, 2013b) and calibration results; the catchments are sorted by mean catchment elevation. F_{LS} is the model efficiency (Eq. 1.1).

No.	River	Area [km^2]	Mean elevation [m a.s.l.]	Regime type [-]	Productivity[a] [-]	Forest [%]	Range F_{LS} [-]
1	Aach	48.5	480	pluvial	0.35	0.33	0.79-0.82
2	Ergolz	261	590	pluvial	0.37	0.41	0.84-0.86
3	Aa	55.6	638	pluvial	0.24	0.22	0.82-0.84
4	Murg	78.9	650	pluvial	0.28	0.31	0.81-0.83
5	Mentue	105	679	pluvial	0.15	0.28	0.79-0.82
6	Broye	392	710	pluvial	0.23	0.25	0.80-0.81
7	Langeten	59.9	766	pluvial	0.35	0.19	0.70-0.74
8	Rietholz	3.3	795	pluvial	0.25	0.21	0.72-0.74
9	Goldach	49.8	833	pluvial	0.11	0.31	0.82-0.83
10	Guerbe	117	873	pluvial	0.30	0.33	0.78-0.80
11	Biber	31.9	1009	pluvial	0.22	0.41	0.82-0.84
12	Kleine Emme	477	1050	nivo-pluvial	0.21	0.35	0.82-0.82
13	Ilfis	188	1051	nivo-pluvial	0.24	0.46	0.78-0.81
14	Sense	352	1068	pluvio-nival	0.23	0.33	0.77-0.79
15	Alp	46.6	1155	nivo-pluvial	0.22	0.45	0.76-0.78
16	Emme	124	1189	nival	0.17	0.32	0.74-0.78
17	Sitter	261	1252	nival	0.31	0.22	0.73-0.74
18	Erlenbach	0.64	1300	nivo-pluvial	0.05	0.60	0.75-0.77
19	Luempenen	0.93	1318	nivo-pluvial	0.50	0.35	0.76-0.77
20	Grande Eau	132	1560	nival	0.20	0.33	0.79-0.81
21	Schaechen	109	1717	nival	0.28	0.16	0.90-0.92
22	Allenbach	28.8	1856	nivo-glaciaire	0.12	0.13	0.73-0.76
23	Riale di Calneggia	24	1996	nivo-pluvial	0.26	0.07	0.80-0.82
24	Dischma	43.3	2372	glacio-nival	0.25	0.02	0.77-0.81

[a] Values of area-weighted catchment average assigned to hydrogeological productivity classes: not=0; little=0.25; variable=0.5; productive=1

Many studies have used scenarios to estimate the impact of climate change on streamflow in general and some that focus on droughts in particular (e.g. Wetherald and Manabe, 1999; Wetherald and Manabe, 2002; Wang, 2005; Lehner et al., 2006). The usual approach is to use simulations of general circulation models or regional climate models (GCM/RCM) with plausible scenarios of greenhouse gas emissions to drive hydrological models. However, there are large uncertainties connected to the GCM and RCM simulations and the choice of bias correction method (Teutschbein and Seibert, 2012; Teutschbein and Seibert, 2013), and the range of resulting impacts is accordingly high. Wilby and Harris (2006) used different GCMs, emission scenarios, downscaling techniques and hydrological model versions to assess uncertainties in climate change impacts and found that the resulting cumulative distribution functions of low flow for the river Thames were most sensitive to uncertainties in climate change scenarios and downscaling. Instead of dealing with these large uncertainties, here we focus on systematic changes. Thus, scenarios that exclude the large sources of uncertainty (climate change scenarios and downscaling) are needed to investigate the different reactions of catchments to droughts.

In this study we address how sensitive different catchments are to meteorological droughts and whether this sensitivity can be linked to a specific type of catchment, classified by catchment characteristics. We aim to answer these questions using a modeling experiment with two different scenarios of increasingly drier meteorological conditions, based on observations.

2 Methods and Data

2.1 Data

We selected 24 Swiss catchments, which vary in size, mean catchment elevation, land cover and geology (Table 1.1). To investigate the main natural underlying processes, only catchments with minor anthropogenic influence were selected, i.e. no catchments with dams, major water extractions or inflow of sewage treatment plants. Additionally, the catchments have, if any, minimal glacier influence and have discharge stations of satisfactory precision during low flow. Daily discharge observations were provided by (FOEN, 2013a). Gridded temperature [°C] and precipitation [mm] data (Frei, 2013) available for Switzerland (MeteoSwiss, 2013) were averaged over each catchment and then used to force the hydrological model. The observation period for discharge data used in this study extended from 1993 to 2012, for the meteorological data from 1975 to 2012. Size, mean catchment elevation, forested land cover, and slope were extracted from the digital elevation map of Switzerland (25 m resolution). A hydrogeological productivity number, which is a measure of hydraulic conductivity and thickness of the aquifer, was derived from the vulnerability map of Switzerland (Spreafico et al., 1992): first, features of the aquifers were classified as productivity: high, variable, low, zero. We assigned a numeric value to each of these productivity classes and computed an area-weighted mean.

2.2 HBV modeling experiment

For the modeling experiment we used the semi-distributed conceptual HBV model (Bergström, 1995; Lindström et al., 1997b) with the version HBV light (Seibert and Vis, 2012). In this study the catchments were separated into elevation zones of 100 m. The model uses different routines (Figure 1.1) to simulate catchment discharge based on time series of daily precipitation and air temperature as well as estimates of long-term monthly potential evapotranspiration.

The routines in HBV include the following:

- Snow routine: snow accumulation and melt are computed by a degree-day method including snow water holding capacity as well as potential refreezing of melt water.
- Soil routine: groundwater recharge and actual evaporation are simulated as functions of the actual water storage in the soil box. The soil moisture storage is called SM.
- Response routine: runoff is computed as a function of water storage in an upper and a lower groundwater box. The groundwater storage (GW) from both groundwater boxes was summed.

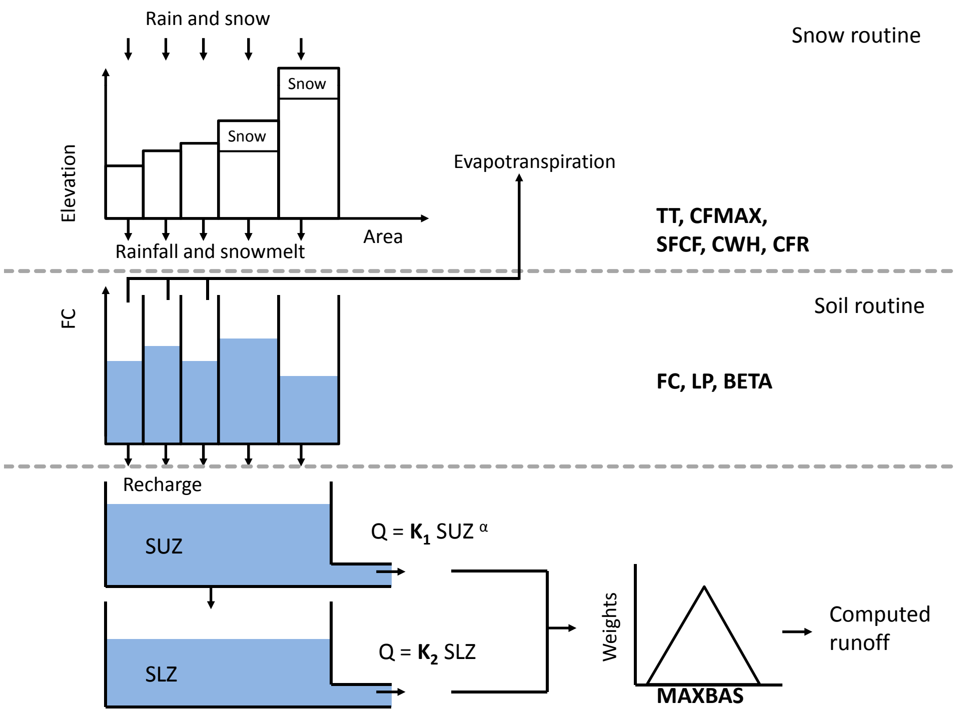

Figure 1.1: Structure of the HBV model.

- Routing routine: a triangular weighting function routes the runoff to the outlet of the catchment.

Detailed descriptions of the model can be found elsewhere (Bergström, 1995; Lindström et al., 1997b; Seibert, 1999). The HBV-light model was calibrated automatically for each of the catchments over the period 1993 to 2012 using a genetic optimization algorithm with subsequent steepest gradient tuning (Seibert, 2000). Parameter uncertainty was addressed by performing 100 calibration trials, which resulted in 100 optimized parameter sets according to a combination of Nash Sutcliffe model efficiency and volume error (F_{LS}, Eq. 1.1 (Lindström et al., 1997b)), where the weighting factor for the latter was set to 0.1, as recommended by Lindström et al. (1997b) and Lindström (1997). F_{LS} ranges between minus infinity for poor fits and 1 for a perfect fit,

$$F_{LS} = 1 - \frac{\sum (Q_{obs} - Q_{sim})^2}{\sum (Q_{obs} - \overline{Q_{obs}})^2} - 0.1 \frac{\sum |(Q_{obs} - Q_{sim})|}{\sum Q_{obs}} \tag{1.1}$$

One simulation was run for each of the parameter sets over the entire meteorological observation period and the simulation results of this ensemble of the 100 selected parameter sets were averaged at each time step to derive the reference simulation. The same procedure was performed for the scenarios.

2.3 Scenario construction

Two precipitation time series were constructed as hypothetical scenarios, over the period 1975 to 2012, with progressively drying conditions:

- Scenario with sorted years (SoYe): All years over the meteorological observation period were sorted from the wettest to the driest year according to the total annual precipitation. Thus, a scenario of modest but continuous progression of drying was constructed.

- Scenario with sorted months (SoMo): For this scenario we shuffled the individual months, with the wettest January together with the wettest February, and so on forming the first year. The second wettest individual calendar months composed the second year. With this approach a scenario was created with a continuous progression of drying in a more extreme manner than SoYe, but still keeping the natural seasonality.

The daily air temperature matching the precipitation from the original time series was re-arranged in parallel to the precipitation scenarios.

2.4 Relative change to long-term conditions

First, we looked at the relative change of each scenario year, x_i, to the long-term mean of the reference simulation, $\overline{x}$.

$$\Delta x_i = \frac{x_i}{\overline{x}} \tag{1.2}$$

where x stands for the variable of interest, and i the year. Δx was calculated for simulated runoff (Q_{sim}), simulated soil moisture storage (SM), and the combined simulated upper and lower groundwater storages ($GW = SUZ + SLZ$) (Fig. 1.1) (Eq. 1.2). Secondly, to assess the catchment sensitivity to the progression of drying we calculated the inter-quartile range (IQR) of Δx. IQR represents the variability during the drying phase and since the scenarios force progressive drying over the course of the years, IQR can be seen as a measure of sensitivity to droughts: the smaller the value of IQR, the less sensitive a catchment is to droughts, and the higher the value of IQR, the more sensitive a catchment is to droughts. This sensitivity of course results from both the local climate variability and modification by specific catchment characteristics. Since the construction of the scenarios was based on annual and monthly precipitation differences, we accounted for the relative influence of the inter-annual variability of precipitation in each catchment on the scenario. For each year the ratio between mean annual precipitation P and long-term mean annual precipitation $\overline{P}$ was calculated (Fig. 1.2). This precipitation ratio was used in the further analysis to account for the potential influence of the inter-annual precipitation variability to enable a comparison between the different catchments. To minimize the influence of the local precipitation variability each IQR was divided by the inter-quartile range of these precipitation ratios (Eq. 1.3). The so modified IQR is referred to as I_{rel}.

$$I_{rel} = \frac{\Delta x_{75} - \Delta x_{25}}{\frac{P}{\overline{P}}_{75} - \frac{P}{\overline{P}}_{25}} \tag{1.3}$$

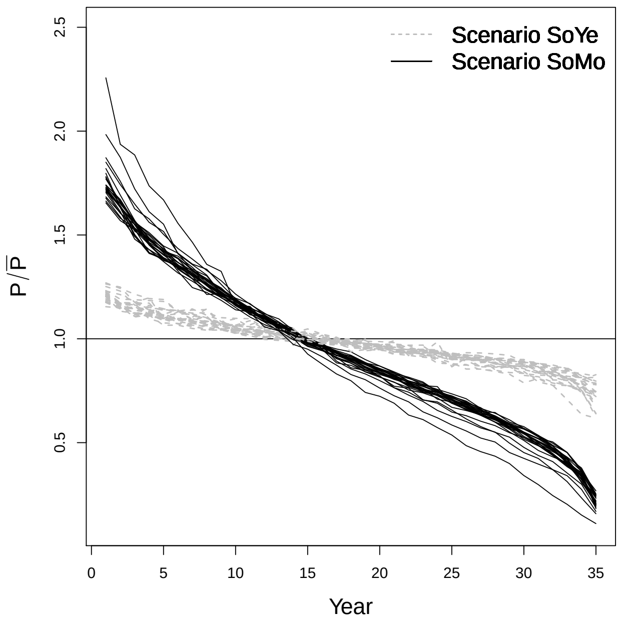

Figure 1.2: Precipitation ratios of the annual precipitations P and the longterm mean precipitation $\bar{P}$ for scenario SoYe and SoMo.

where Δx_{75} is the 75^{th} percentile of Δx and Δx_{25} the 25^{th} percentile of Δx. Even though this I_{rel} includes both wet and dry years, it gives an overall impression of the reaction of a catchment to the progression of drying. We also compared the extreme end of each scenario (driest year of both scenarios) with the long-term mean to account for drought more specifically. The extreme end of each scenario was additionally compared to the driest year from the reference simulation in order to determine in which seasons the strongest effect of drying was found. Further, to find catchment controls on the sensitivity of catchments to droughts, we analyzed the correlations between specific catchment characteristics (Table 1.1) and sensitivities. The correlations were calculated using the Spearman rank correlation to detect rank correlations between catchment characteristics and sensitivities. The significance of the correlations was evaluated using the p-value of the distributions, where correlations with a p-value of <0.05 are considered significant.

To target drought characteristics more specifically, we counted the days per year that exceeded the 90^{th} streamflow percentile (Q_{90}) of the respective reference simulation. Q_{90} is a commonly used threshold value to define hydrological drought periods. Again, we calculated a relative change (Eq. 1.2), here with the exceedance days of Q_{90} as x. Other indices describing the influence of the progression of drying at its extreme dry end, are the ratios of the mean of the driest year of each scenario and the long-term mean ($\Delta Q_{DriestSoYe}$ for scenario SoYe; $\Delta Q_{DriestSoMo}$ for scenario SoMo). The smaller these indices are, the more sensitive the respective catchments are to droughts.

Further, we studied the summer of 2003 as one of the historical droughts that falls in the observation period of this study in another simulation experiment making use of the scenario SoMo. Here, we used the last years of the scenario SoMo up to the end of May followed by the actual series of summer 2003 starting from 1 June. In this way for each catchment we simulated how much more the catchment would have been affected if the preceding months to the 2003 drought event would have been drier than in the actual observation. For all catchments a further index was calculated describing the sensitivity of the catchments to drier initial conditions and thus also to droughts by dividing the mean of the SoMo scenario based simulation with the drier initial conditions for the summer months of 2003 (June-August) by the mean of the reference simulation for the same months. This index was called ΔQ_{2003} for $Qsim$, ΔSM_{2003} for SM, and ΔGW_{2003} for GW. The smaller these indices are, the less sensitive the respective catchments are to droughts.

3 Results

3.1 Inter-annual variation

All catchments could be calibrated satisfactorily with median F_{LS} values (Eq. 1.1) ranging between 0.73 and 0.92 (Table 1.1). The relative change of the different variables clearly indicated a progression of drying of streamflow as well as of the storages, where the relative change of the continuous drying for all catchments was smallest for SM for both scenarios (Figure 1.3).

The SoMo scenario generally resulted in stronger responses to the drying and the relative changes specific for the different catchments became more pronounced than in scenario SoYe. The higher elevation catchments compared to the lower elevation catchments show larger Δ_{Qsim} values for the wet years at the beginning of the drying scenario. However, for the dry years of the drying scenario the Δ_{Qsim} values are smaller for the high elevation catchments than for the low elevation catchments. During drier conditions than the long-term mean, the Δ_{Qsim} values of the catchments with higher elevations were smaller compared to lower elevation catchments. The same can also be seen for Δ_{GW} in the scenario SoMo where the change from wetter (above 1) to drier (below 1) relative to the longterm GW mean shows more variability between the catchments than for Δ_{Qsim} (Figure 1.3).

The general behavior was shown using four catchments as examples comparing the long-term mean and the driest year of the reference simulation (Figure 1.4). Scenario SoYe resulted, most of the time, in streamflow values below the long-term mean. However, the scenario did not always result in lower streamflow values compared to the long-term mean, but had rather seasons with pronounced lower flows: this was the case in fall/winter as well as in summer, where the hydrograph of the SoYe scenario falls entirely below the long-term mean for the pluvial Mentue catchment. For the nival catchments (Ilfis, Sitter, Emme) the hydrograph from SoYe was below the long-term mean streamflow during the spring flood as well as during late summer. However, the difference between long-term mean streamflow and the streamflow from scenario SoYe varied remarkably between the catchments although the timing of the pronounced lower flows appeared to occur simultaneously. The overall difference between the long-term mean and the scenario SoYe, which can be seen in the cumulative sums, confirms this variation between the catchments and thus their sensitivity to the continuous drying (Figure 1.4). The difference between the last year SoYe and the driest year of the reference simulation resulted from the different initial conditions given by the preceding summer and was

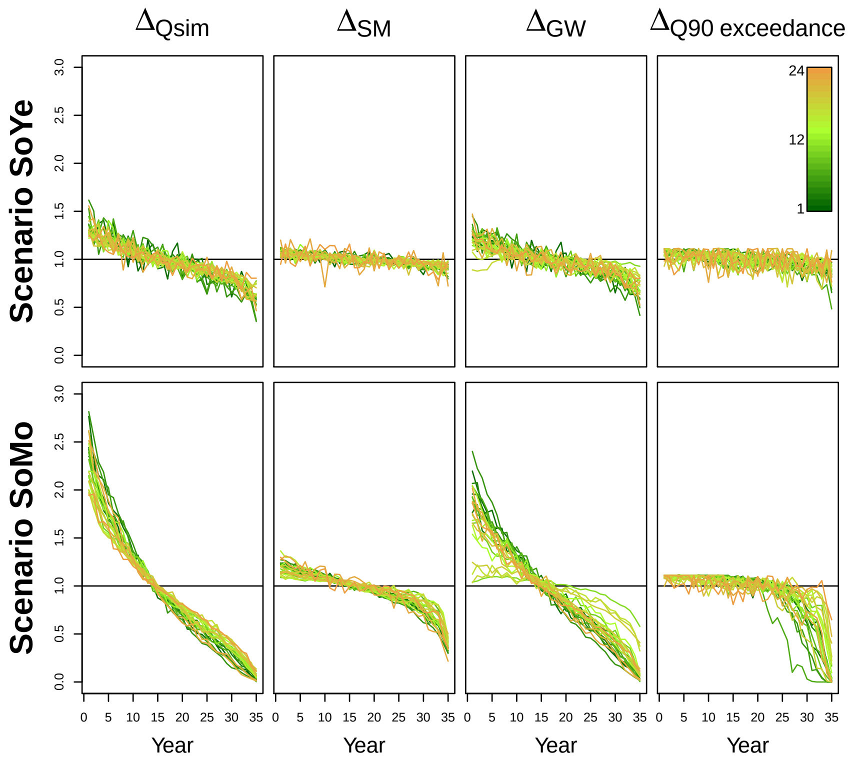

Figure 1.3: Relative change of Q_{sim}, SM and GW and the Q_{90} exceedance days for the two scenarios to the longterm reference for all catchments. Each color stands for one catchment number (Table 1.1), where the greener colors indicate catchments at lower mean elevation and the more brownish colors were used for catchments at higher mean elevation.

very small. The driest year of scenario SoMo resulted always in streamflow values below the long-term mean as well as below the driest year of the reference simulation and of the SoYe scenario for all catchments. For the pluvial Mentue catchment discharge run dry for the driest year of the scenario. For the catchments with some snow influence there are periods of higher streamflow in spring and summer, however with a very reduced spring flood as compared to the SoYe scenario or the longterm mean. For the scenario SoMo, the cumulative sums show that the annual difference between long-term mean and the scenario varies for the different catchments.

3.2 Low flow frequency

The relative difference of the frequency of days that were exceeding the Q_{90} threshold was small for the SoYe scenario (Figure 1.3). Even though over the course of the years a slight decrease of days exceeding the threshold could be noticed, there were still years at the end of the scenario that had longer durations of exceeding days than the longterm mean. Note, that the upper boundary is given by the fixed number of days that are exceeding Q_{90} per year and the maximum days per year. For the SoMo scenario, however, there is a strong decrease in days exceeding the Q_{90} threshold with the progression of drying. In

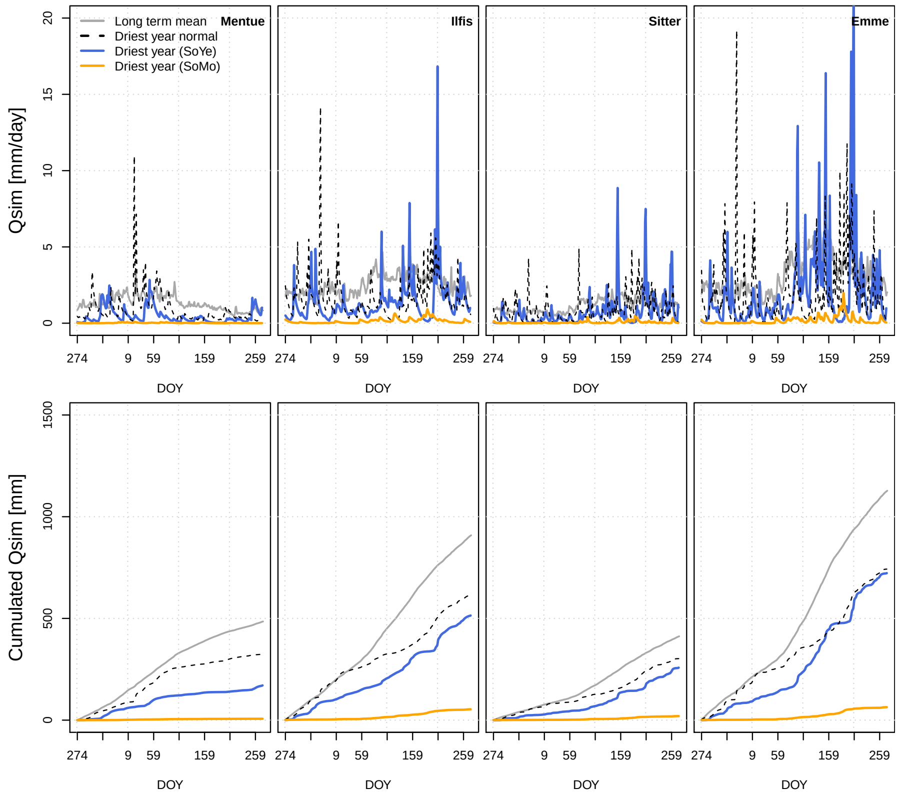

Figure 1.4: Q_{sim} and cumulated Q_{sim} for long-term mean, driest year of the reference simulation as well as the driest years of the two scenarios for four example catchments.

this scenario a difference between the catchments also became apparent: in the relatively wetter years, the lower elevation catchments already start to have less days above the threshold, i.e. are more vulnerable to droughts. In the medium dry years of the scenario the higher elevation catchments also show less days above the threshold compared to the long-term mean. The highest elevation catchments follow in even drier years of the scenario to show less days above the threshold compared to the long-term mean. In comparison to scenario SoYe, in scenario SoMo, the highest elevation catchments show a clear decrease in days above the Q_{90} threshold at the dry end of the scenario.

The historical drought event of the summer 2003 and how it would have changed with different initial conditions for the different catchments is shown for the four example catchments (Figure 1.5). While for the Mentue, Ilfis and Sitter catchments the influence of the drier initial conditions can be seen relatively long into the summer months, for the Emme catchment, this memory is comparably short. However, looking at the storages SM and GW for the reference simulation as well as the simulation with drier initial conditions shows that the causes for longer or shorter influence are not the same for the different catchments: the important storage for the effect of the initial conditions for Mentue and Ilfis is composed of both storages, while for the Sitter and the Emme catchments SM seems to be stronger and important for longer than GW.

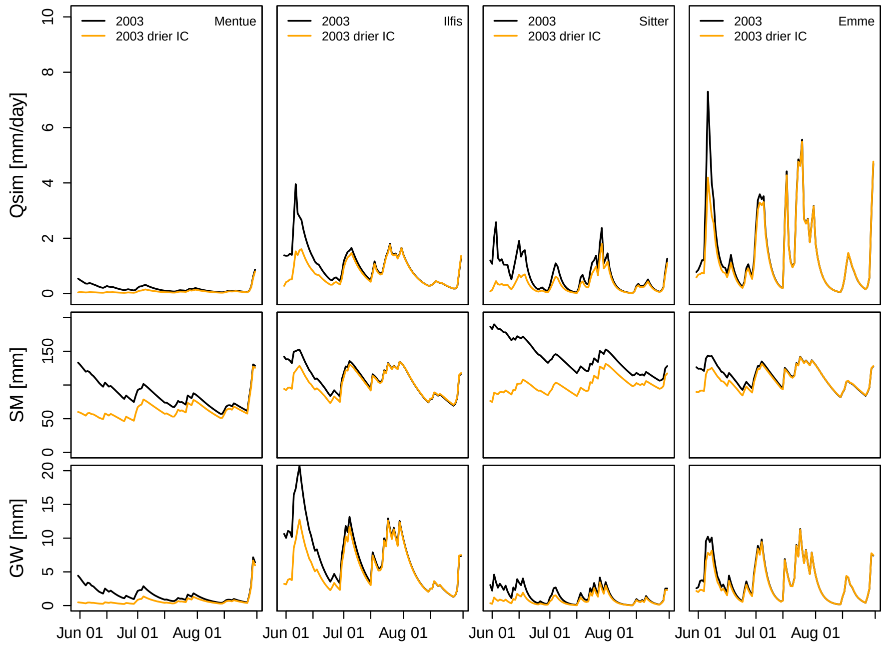

Figure 1.5: Simulation (median of 100 simulations) of the summer drought 2003, original and with drier initial conditions (IC).

3.3 Importance of catchment characteristics

The values of I_{rel} (Q_{sim}) were significantly correlated with catchment mean elevation, size and slope, respectively (Figure 1.6). Mean catchment elevation and drought sensitivity were correlated with higher mean catchment elevations related to lower drought sensitivities. Steeper slopes are also related to lower drought sensitivities. Even though there was a significant correlation between mean catchment elevation and slope, the highest elevation catchments do not always have the steepest slopes. Hence, it makes sense to look at both slope and mean catchment elevation individually. For SM the I_{rel} values were significantly correlated with size and slope, while for GW the I_{rel} values were correlated with mean catchment elevation and slope. The variables describing hydrogeology (productivity number) as well as land cover (percentage forested area) had no significant influence on the sensitivity of the catchments studied to droughts.

A summary of all indices can be found in Table 1.2. The drought targeting indices (IQR of days above the threshold Q_{90}, $\Delta Q_{DriestSoYe}$, $\Delta Q_{DriestSoMo}$, and changes of summer 2003 with drier initial conditions ΔQ_{2003}, ΔSM_{2003} and ΔGW_{2003}) could also be related to the catchment characteristics (Figure 1.7); most of them were correlated with size, elevation or slope of the catchment: IQR of days above the threshold, Q_{90} as well as ΔQ_{2003}, were significantly correlated with size and slope of the catchment. The ratios of the driest years of the two scenarios $\Delta Q_{DriestSoYe}$ and $\Delta Q_{DriestSoMo}$ were significantly correlated with size and elevation, respectively. ΔSM_{2003} was correlated with mean elevation, slope and size of the catchment.

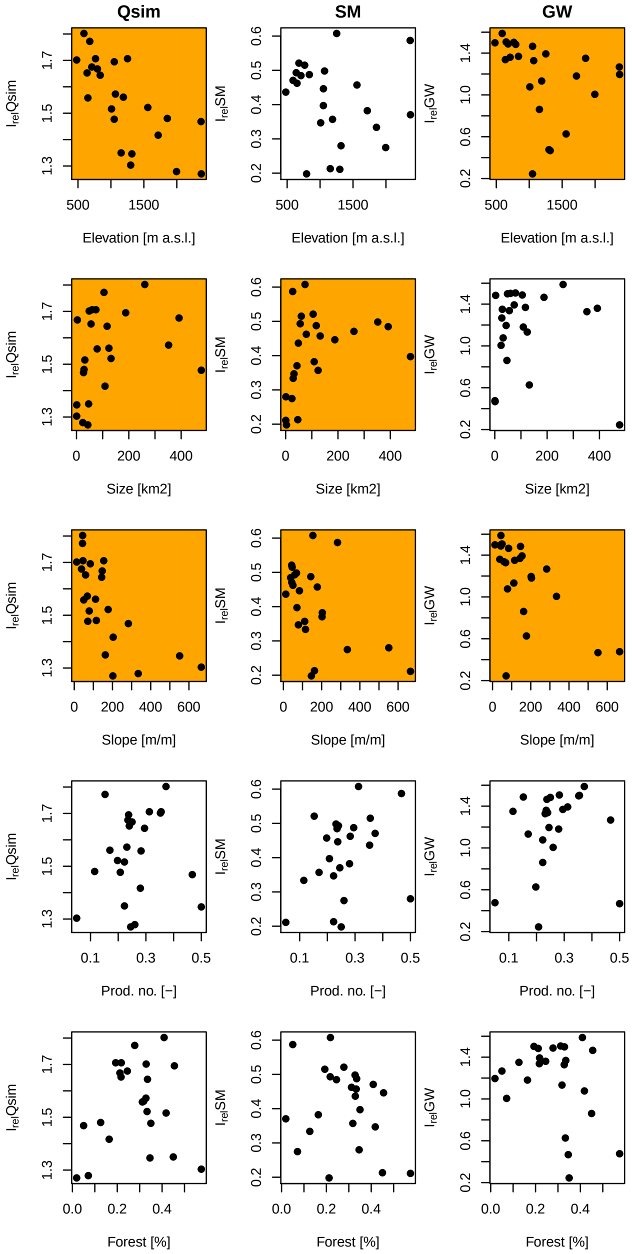

Figure 1.6: I_{rel} for Q_{sim}, SM, and GW compared to simple catchment characteristics. The orange background indicates a significant correlation (5% level) between the respective I_{rel} and catchment characteristic. For an explanation of the productivity number (prod.no.) see description in section Methods and Data.

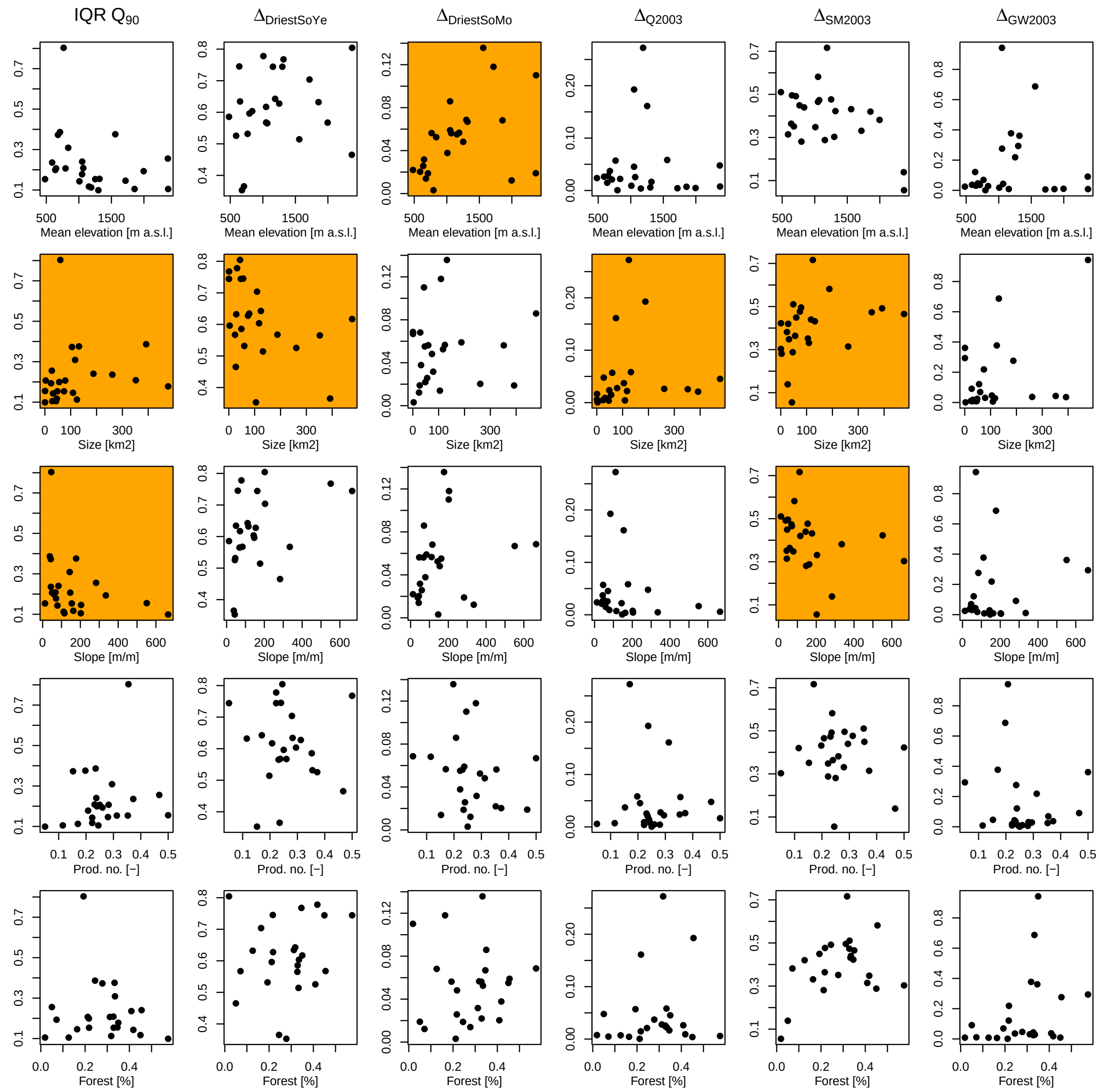

Figure 1.7: Indicators to drought sensitivity: days above the threshold Q_{90}, $\Delta Q_{DriestSoYe}$, $\Delta Q_{DriestSoMo}$, and changes of summer 2003 with drier initial conditions ΔQ_{2003}, ΔSM_{2003} and ΔGW_{2003} compared to simple catchment characteristics. The orange background indicates a significant correlation (5% level) between the respective indicator and catchment characteristic. For an explanation of the productivity number (prod.no.) see description in section Methods and Data.

4 Discussion

In the analysis of the simulations from the scenarios, which clearly depend on the inter-annual variability of precipitation for each catchment, we removed the effect of precipitation variability by dividing the IQR values by the inter-quartile range of the precipitation ratio. Following many studies that document the sensitivity of streamflow to climate and climate change, Schaake and Waggoner (1990), Dooge (1992), and Sankarasubramanian et al. (2001) introduced and applied the so called streamflow elasticity, which describes the sensitivity of streamflow to precipitation. The streamflow elasticity was developed as a robust, unbiased approach that on average and over many applications might discern the true sensitivity of streamflow to climate (Sankarasubramanian et al., 2001). Similar to our approach, the streamflow elasticity is calculated by taking annual streamflow and precipitation into account (Sawicz et al., 2011). For the comparability of the sensitivity to the drying in our approach instead we ensured that the streamflow elasticity did not influence the results of the scenario simulations, i.e. that the weather difference between the different years related to each catchment should not overprint the catchment-specific properties.
The scenarios were constructed by applying sorted annual or monthly precipitation, while air temperature was not considered explicitly. For example Null et al. (2010) considered air temperature and analyzed streamflow and particular low flow sensitivities to climate change by using scenarios with increased temperatures, but constant precipitation for mountain catchments. However, the results of previous case studies considering total streamflow response to changes in precipitation and temperature indicated that future total streamflow is more sensitive to precipitation than to temperature (Lettenmaier et al., 1999; Nijssen et al., 2001).
Another issue related to the construction of our scenarios is that the preceding wetness of the season was not considered while sorting. This could lead to actual drier or wetter initial conditions for the following year than indicated by the annual sum, particularly for the SoYe scenario. We tried to minimize this effect by using hydrological years starting on October 1, and not calendar years. Still, there could have been a dry summer in an otherwise relatively wet year which then serves as initial conditions for the following year. However the effect should be low compared to a start in winter with, for instance, a large snow cover at the end of an otherwise dry year.
We looked at the effects of the continuous progression of drying on the different catchments and found that, in general, even modest drying led to a continuous reduction of streamflow, soil moisture and groundwater storage on the one hand and on the other hand the moderate scenario already revealed catchments that were more sensitive to droughts than others. With the more extreme scenario the picture became even clearer. However, for the drought characteristic duration of days above the Q_{90} threshold, an effect was only visible after applying the more extreme scenario. The driest year of the moderate scenario showed seasons with lower than the long-term mean streamflow values, that differed for catchments with different streamflow regimes. In the higher elevation catchments the snow component needed to be considered in summer and fall, opposite to the long dry summer and fall seen for lower elevation catchments. Hence, there were again higher streamflow values visible in late summer in the higher elevation catchments. This could be explained by a filling of the storages in spring with snow melt

water, that kept the storages at a higher level than would be possible if only rainfed (at least in the temperate humid climate of Switzerland). Further differences between the catchments with nival regimes have then to be accounted for by different storage release characteristics. This could be confirmed by the analysis of the historical drought in the summer of 2003 compared to a scenario with drier initial conditions since the storages for the different catchments contributed in different proportions to the reduced streamflow under drier initial conditions.

The relative differences between the catchments might appear small however, they are not minor. It is interesting that the initial conditions can have noticeable impacts even when looking at a whole year. The differences due to initial conditions varied between about 50% and 80%, which is in the same order of magnitude as of what might be expected due to climate change (Nijssen et al., 2001).

Comparing the drought sensitivities to catchment characteristics revealed that for both streamflow (I_{rel}) as well as duration of days above the Q_{90} threshold mean catchment elevation, size and slope were the main controls. Kroll et al. (2004), who tested different catchment characteristics as to their suitability to improve the regionalization of low flows in the US, found that signatures describing hydrogeology, slope and size and also elevation were important. However, while size was an important predictor for almost every region they investigated, elevation improved low flow prediction only in a few regions of the US. For soil moisture storage only size and slope control drought sensitivity and for groundwater storage only elevation and slope control drought sensitivity. This means that the variability of the storages is not controlled by the same catchment characteristics as the resulting streamflow. However, streamflow as it integrates the catchment processes, showed all the controls of the storages. The fact that mean catchment elevation is important for drought sensitivity in streamflow can be partly explained by snow in higher elevations. Other reasons like larger subsurface storage features in higher elevation catchments are indicated by the relationship between simulated groundwater storage and mean catchment elevation.

The catchment characteristic hydrogeology could be expected to be correlated to a storage dependent drought sensitivity (Stahl and Demuth, 1999; Kroll et al., 2004), however we could not find any relationship. It could be that the hydrogeological productivity number was not an appropriate measure for storage and release. It could also be that the other controls dominated and hence secondary effects like geology or land-use, which are also very diverse and show a high variability among the catchments, did not show any correlation.

The results that are derived from the modeling experiment contain potential sources of uncertainty, i.e. mainly the choice of the hydrological model and its associated structure and parametrization. The uncertainty from the model parametrization was addressed by an ensemble approach, which generated a more robust simulation than would have been the case for single “best” parametrization. Concerning the model structure we can assume that the main indication of the results of the streamflow simulation should be similar for different conceptual hydrological models, whereas we can expect some differences in the simulated storages.

Clearly it must be stated that the scenarios that were used did not aim to be realistic. For instance, the precipitation in the scenarios decreased intentionally over the course of the years, which causes unnatural autocorrelations. Other studies that use, e.g., GCM output extreme climate change scenarios for climate impact studies, keep the natural variation of precipitation from year to year (e.g., Miller et al., 2003; Burke et al., 2006). Instead the scenarios in this study were constructed to get an idea of how strongly a catchment would react to a moderate and to an extreme progression of drying in comparison with a sample of other catchments from the temperate humid climate of Switzerland. The scenarios were also derived in order to better understand how strongly initial conditions affect hydrological droughts, and were appropriately constructed for this purpose.

As a next step it would be interesting to perform an analysis similar to the one in this study for other regions as well as to find a system of general drivers that make a specific catchments vulnerable to droughts or not. A ranking for the different catchments that could help drought managers as a starting point to decide on which catchments are more vulnerable to droughts can easily be derived from our results. In addition to the scenarios used in this study, there is also the possibility to construct scenarios that have time fractions for sorting that are in between the yearly and the monthly construction of this study, for example, scenarios using half a year, a quarter of a year or two months.

5 Conclusions

This study demonstrates that hypothetical scenarios can be used to evaluate the sensitivity of catchments to droughts. The reaction of streamflow as well as soil moisture and groundwater storages to a continuous progression of drying was analyzed both in general as well as focused on drought characteristics and on one historical drought event. Our analysis showed that mean catchment elevation, size and slope were the main controls on the sensitivity of the catchments to drought. The results suggest that higher elevation catchments with steeper slopes were less sensitive to droughts than lower elevation catchments with less steep slopes. The soil moisture storage was significantly correlated to catchment size, where we found smaller catchments to be less sensitive to droughts than larger catchments. We did not find a clear connection between drought sensitivity and hydrogeology. Generally, for water resource management it is important to look at both streamflow sensitivity and storage sensitivity to droughts. With our model-based approach the sensitivity of both can be easily estimated. This approach can serve as a starting point for water resources managers to understand the vulnerability of their catchments.

Table 1.2: Drought indicators for all catchments. Note that for all indicators the smaller the value the less sensitive the catchment is to drying, and the smaller the values of $\Delta Q_{driestSoYe}$ and $\Delta Q_{driestSoMo}$, the more sensitive the catchment is to drying.

Catchment	$I_{relQsim}$	I_{relSM}	I_{relGW}	IQR_{Q90}	ΔQ_{2003}	ΔSM_{2003}	ΔGW_{2003}	$\Delta Q_{driestSoYe}$	$\Delta Q_{driestSoMo}$
Aach	1.701	0.436	1.499	0.154	0.024	0.510	0.025	0.586	0.022
Ergolz	1.802	0.471	1.588	0.236	0.026	0.314	0.037	0.526	0.020
Aa	1.653	0.493	1.338	0.199	0.015	0.364	0.121	0.745	0.026
Murg	1.558	0.462	1.507	0.207	0.028	0.496	0.031	0.634	0.032
Mentue	1.772	0.521	1.487	0.373	0.037	0.351	0.046	0.353	0.014
Broye	1.675	0.485	1.360	0.386	0.021	0.492	0.036	0.365	0.019
Langeten	1.706	0.515	1.503	0.803	0.057	0.449	0.069	0.532	0.056
Rietholz	1.668	0.198	1.483	0.207	0.001	0.281	0.001	0.596	0.003
Guerbe	1.644	0.488	1.369	0.309	0.022	0.440	0.028	0.604	0.053
Biber	1.516	0.347	1.077	0.143	0.009	0.348	0.018	0.778	0.038
Kleine Emme	1.477	0.397	0.245	0.178	0.045	0.465	0.943	0.617	0.086
Ilfis	1.695	0.446	1.465	0.240	0.193	0.582	0.276	0.568	0.059
Sense	1.572	0.498	1.328	0.208	0.026	0.473	0.043	0.565	0.056
Alp	1.350	0.213	0.861	0.117	0.004	0.288	0.009	0.744	0.055
Emme	1.561	0.357	1.133	0.113	0.272	0.716	0.377	0.643	0.057
Sitter	1.706	0.608	1.392	0.154	0.161	0.477	0.219	0.628	0.048
Erlenbach	1.303	0.211	0.476	0.099	0.006	0.303	0.294	0.744	0.069
Luempenen	1.346	0.280	0.467	0.155	0.017	0.423	0.361	0.768	0.067
Grande Eau	1.522	0.457	0.626	0.376	0.058	0.431	0.687	0.514	0.136
Schaechen	1.417	0.382	1.181	0.146	0.004	0.331	0.006	0.704	0.118
Allenbach	1.480	0.334	1.350	0.105	0.007	0.420	0.008	0.632	0.068
Riale di Calneggia	1.279	0.275	1.005	0.193	0.005	0.382	0.011	0.567	0.012
Ova da Cluozza	1.468	0.587	1.267	0.256	0.048	0.139	0.091	0.465	0.019
Dischma	1.270	0.370	1.196	0.105	0.008	0.055	0.008	0.804	0.110

APPENDIX 2

Predictability of low flow - an assessment with simulation experiments*

Abstract

Since the extreme summer of 2003 the importance of early drought warning has become increasingly recognized even in water-rich countries such as Switzerland. Spring 2011 illustrated drought conditions in Switzerland again, which are expected to become more frequent in the future. Two fundamental questions related to drought early warning are: 1) How long before a hydrological drought occurs can it be predicted? 2) How long are initial conditions important for streamflow simulations? To address these questions, we assessed the relative importance of the current hydrological state and weather during the prediction period. Ensemble streamflow prediction (ESP) and reverse ESP (ESP_{rev}) experiments were performed with the conceptual catchment model, HBV, for 21 Swiss catchments. The relative importance of the initial hydrological state and weather during the prediction period was evaluated by comparing the simulations of both experiments to a common reference simulation. To further distinguish between effects of weather and catchment properties, a catchment relaxation time was calculated using temporally constant average meteorological input. The relative importance of the initial conditions varied with the start of the simulation. The maximum detectable influences of initial conditions ranged from 50 days to at least a year. Drier initial conditions of soil moisture and groundwater as well as more initial snow resulted in longer influences of initial conditions. The catchment relaxation varied seasonally for higher elevation catchments, but remained constant for lower catchments, which indicates the importance of snow for streamflow predictability. Longer persistence seemed to also stem from larger groundwater storages in mountainous catchments, which may motivate a reconsideration of the sensitivity of these catchments to low flows in a changing climate.

*Accepted for publication in the Journal of Hydrology as: Staudinger, M. and Seibert, J., Predictability of low flow - an assessment with simulation experiments.

1 Introduction

In many parts of the world people are aware of droughts as natural hazards with significant impacts on many sectors especially when they persist for long periods or occur frequently (e.g. Tallaksen and van Lanen, 2004; Dijk et al., 2013; Viste et al., 2013). However, only recently, scientists and stakeholders in Europe have become concerned not only about floods and their forecasting, but also about droughts. Drivers of this increasing interest include recent droughts such as in summer 2003 (Rebetez et al., 2006) and in spring 2011, which have made water rich countries like Switzerland become more aware of impacts and risks related to droughts. So far, the main concerns in Europe regarding droughts are of economic, environmental, and social importance (e.g. Stahl et al., 2012). During and after droughts, conflicts between different water users can become more frequent and water management has to adapt to meet the different interests as well as possible. For these reasons drought early recognition has become an issue. The basic objective of drought early recognition is to provide timely warning, so that damages can be reduced or even avoided. However, little has been done regarding forecasting and early warning of droughts in Europe. The severity of a drought depends clearly on the climatological deficit of water, but also on the hydrological system that has to cope with this deficit. There were many attempts to quantify droughts by indices based on meteorological variables such as the Palmer drought severity index (Palmer, 1965), deciles (Gibbs and Maher, 1967), the surface water supply index (Shafer and Dezman, 1982), the standardized precipitation index (McKee et al., 1993) or the standardized precipitation and evapotranspiration index (Vicente-Serrano et al., 2010). Each of these indices has its own strengths and weaknesses. Drought indices based on meteorological variables are important, but not sufficient to describe and understand the severity of a hydrological drought. Hence, to recognize locally critical conditions early and provide that information to decision makers, requires both information of the climatological anomalies as well as an understanding of the underlying hydrological systems.

The persistence of a system is a measure of how a hydrological condition at a certain point in time can influence the following period and can also be seen as the memory of the system. Catchments with a small storage also usually have a small persistence while catchments with large storages can have longer persistences. The predictability of streamflow and other hydrological variables is highly connected to persistence and there exist various methods to estimate persistences. A classical approach to estimate short term persistence is to calculate the autocorrelation of the time series of streamflow observations (e.g. Vogel et al., 1998; Pagano and Garen, 2005). Applying the autocorrelation to highly seasonal data like streamflow data means that they first need to be de-seasonalized before a signal other than seasonality can be found from the autocorrelation can be found. De-seasonalization procedures for hydrological data, however, often require calibration themselves, as the seasonality rarely corresponds to calendar dates (Hipel and McLeod, 1994).

Several recent studies try to quantify the impact of initial conditions on the predictability of hydrological conditions. Snow cover (Gobena and Gan, 2010; Mahanama et al., 2012), catchment size (Li et al., 2009), North Atlantic Oscillation (NAO), El Niño-Southern Oscillation (ENSO) driven by the Sea Surface Temperature (SST) (e.g. Bierkens and Van Beek, 2009) are generally found to be sources of predictability and they are all

highly dependent on the region, system and season. While temperature and precipitation are in part predictable because of the low-frequency variability in global energy stores, particularly in the ocean, (Westra and Sharma, 2010; Feng et al., 2011), on a local scale there are feedbacks because of, for instance, albedo or catchment moisture storages that affect the partitioning between sensible and latent heat fluxes. Predictability in streamflow is controlled by storages, including snow, soil moisture and groundwater, which attenuate the high-frequency rainfall variability to a lower-frequency streamflow response. Singla et al. (2012) assessed the predictive skill of seasonal hydrological forecast in France with two experiments looking at the influence of land surface initial states on the one hand and atmospheric forcing on the other hand. They focused on the spring season as it is critical to the onset of low flows and droughts. One of their important findings was that the predictability of hydrological variables in France mainly depends on temperature and precipitation in lower elevation areas and mainly on snow cover in high mountains. We built on these studies by looking at the predictability of streamflow with focus on low flows in Switzerland using a conceptual hydrological model. These models are important tools in hydrology as they are able to capture dominant catchment dynamics while remaining parsimonious and computationally efficient (Kavetski and Kuczera, 2007). Conceptual hydrological models can reach, for specific purposes, considerable performance and, thanks to their computational efficiency, can also be used in ensemble prediction systems (Cloke and Pappenberger, 2009). In flood forecasting systems conceptual models like the NAM model (Van Kalken et al., 2004), the Sacramento model (Grijsen et al., 1993), the PDM model (Moore and Jones, 1997) and the HBV model (Bürgi, 2002) are often applied and use for low flow ensemble forecasting is also emerging (Fundel et al., 2013).
In this study we used the HBV model (Bergström, 1992; Lindström et al., 1997b) to perform streamflow simulation experiments and to answer the following questions: How long is the persistence of the initial hydrological state in model simulations of streamflow and does it vary in space and time? Can the persistence be attributed to catchment storage?

2 Data and Methods

2.1 Data

The catchments investigated in this study are meso-scale (3 to 350 km^2), near natural catchments located in Switzerland (Figure 2.1). The mean elevation of the catchments ranges between 480 m a.s.l. and 2400 m a.s.l. (Table 2.1). Henceforth, specific catchments are referred to by catchment numbers (Table 2.1). The data used are daily streamflow from the selected Swiss catchments over the period 1970 to 2008 (FOEN, 2011). The meteorological forcing variables for the HBV model, precipitation and temperature, stem from interpolated observations from climate stations (MeteoSwiss) in Switzerland. The selection of the meteorological stations as well as interpolation and aggregation of the variables for each catchment were carried out by the pre-processing tool WINMET (Viviroli et al., 2009). In brief, the spatial and temporal interpolation of observed meteorological variables was based on elevation-dependent regression, inverse distance weighting, Kriging and a simple elevation lapse-rate for temperature data (more

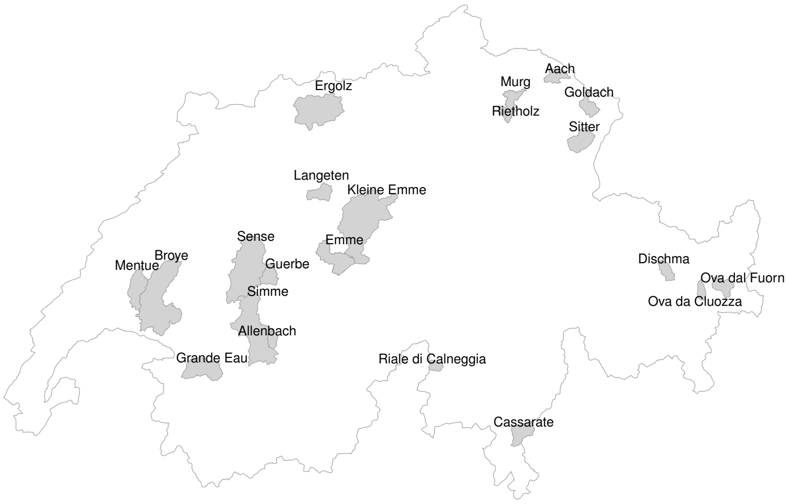

Figure 2.1: Location of the selected Swiss catchments.

details can be found in Viviroli et al. (2009)). A clear seasonal variation of precipitation can be observed for the catchments included here, with winter months receiving about half of the precipitation compared to summer months. The inter-annual variation is similar for all months and about twice as large as the seasonal variation.

2.2 Methods

To quantify the persistence of current hydrological states in streamflow and the influence of weather during prediction we set up three model experiments using the hydrological model HBV in the version HBVlight (Seibert and Vis, 2012) (Figure 2.3).

2.2.1 Model calibration

The HBV model was calibrated for each catchment with the genetic calibration algorithm (GAP), which is included in HBVlight (Seibert and Vis, 2012). With GAP, optimized parameter sets are found by an evolution of parameter sets using selection and recombination (Seibert, 2000). An ensemble of 100 parameter sets was generated for each catchment, based on 100 calibration trials. The mean absolute relative error, F_{MARE} (eq. 2.1), served as the objective function for the calibration, as the emphasis was on low to medium flows. Its values range between minus infinity and the optimum at one.

$$F_{MARE} = 1 - \frac{1}{n}\sum_{i=1}^{n}\frac{|Q_{obs}(i) - Q_{sim}(i)|}{Q_{obs}(i)} \tag{2.1}$$

2.2.2 Estimation of persistence and catchment relaxation

The model input consists of time series of daily precipitation and temperature as well as mean monthly potential evapotranspiration (Penman, 1948). The first two experiments a) and b) were set up much like the experiments in the study of Shukla and Lettenmaier (2011) and the ensemble streamflow prediction (ESP) and the reverse ensemble

Table 2.1: Catchment properties (FOEN, 2011).

Number	Catchment	Area [km^2]	Mean elevation [m a.s.l.]	Regime type	Pores [%]	Fissures [%]	Karst [%]
1	Aach	48.5	480	pluvial	14.9	85.1	0.0
2	Ergolz	261	590	pluvial	10.0	0.0	90.0
3	Murg	78.9	650	pluvial	35.0	65.0	0.0
4	Mentue	105	679	pluvial	77.6	22.4	0.0
5	Broye	392	710	pluvial	65.0	35.0	0.0
6	Langeten	59.9	766	nivo-pluvial	18.3	81.7	0.0
7	Rietholz	3.3	795	nivo-pluvial	0.0	100	0.0
8	Goldach	49.8	833	nival	24.3	75.7	0.0
9	Cassarate	73.9	990	pluvial	0.0	100.0	0.0
10	Sitter	261	1040	pluvial	27.7	39.0	33.3
11	Guerbe	117	1044	nivo-pluvial	77.0	17.7	5.3
12	Kleine Emme	477	1050	nivo-pluvial	50.0	35.0	15.0
13	Sense	352	1068	pluvio-nival	36.7	56.8	6.5
14	Emme	124	1189	nival	12.5	86.5	1.0
15	Grande Eau	132	1560	nival	0.0	60.0	40.0
16	Simme	344	1640	glacio-nival	0.0	75.0	25.0
17	Allenbach	28.8	1856	nivo-glaciaire	48.0	44.5	7.5
18	Riale di Calneggia	24	1996	nivo-pluvial	18.6	81.4	0.0
19	Ova dal Fuorn	55.3	2331	glacio-nival	6.3	14.7	74.0
20	Ova da Cluozza	26.9	2368	glacio-nival	21.3	1.0	77.7
21	Dischma	43.3	2372	glacio-nival	31.2	68.8	0.0

streamflow prediction (ESP_{rev}) approach of Wood and Lettenmaier (2008) (Figure 2.2). However, in this study 100 parameterizations were used for each ensemble member, which allows more robust interpretation by using the ensemble mean as well as quantification of parameter uncertainty effects. Experiments a) and b) evaluate both the influence of initial conditions and weather during prediction on the prediction skill.

The simulation experiments differed in the time series that were used as warming up periods to derive initial conditions, and the time series that were used during the prediction period. In experiment a), during the warming up phase the HBV model was forced with different meteorological time series and the forcing during the prediction was the climatology for all simulations. The climatology, i.e., the long term mean annual series of precipitation and temperature, was computed as 365 arithmetic means of the different years. Experiment b) was the reversed version of experiment a); the time series had identical initial conditions, stemming from the climatology. In the simulations (365 days each), the HBV model was forced with different meteorological time series to derive 'predictions' (Figure 2.2). For both experiments reference runs were performed: in experiment a) the long term mean was used for both warming up and simulation, in experiment b) the same year as in the experimental run was used for the simulation and the previous chronological year was used for the warming up period. By comparison to reference simulations, the two experiments can serve to estimate streamflow persistences that can again be an estimate of the potential streamflow predictability.

A third experiment was designed to distinguish further between the influence of the catchments themselves and the meteorological conditions. A relaxation time for the catchments was calculated, defined as the time needed for the system to reach a new

equilibrium after being brought off balance (e.g. Graf, 1977; Ahnert, 1987; Roering et al., 2001). The warming up in experiment c) was the same as in experiment a). The forcing during the simulation was kept constant and the average annual daily precipitation, mean annual temperature and zero evapotranspiration were used. The precipitation was then distributed to correspond to realistic conditions with precipitation on about 30% of the days, i.e., three times the average precipitation was used as forcing on every third day and zero precipitation otherwise. Before running the simulations the initial snow conditions were all set to zero. This was done to remove the influence the melting of accumulated snow had on the relaxation time estimation, which would obviously have lead to longer relaxation times for catchments with large snow storage. Hence, the catchment relaxation time in this study is the streamflow persistence under constant forcing of temperature and periodic forcing of precipitation.
We defined the persistence [days] in the simulated streamflow as the period from the start of the experiment simulation to the point of convergence (absolute average difference equal to 0.002 mm/day) to the respective reference simulation. After convergence there is no impact of the initial conditions visible in the simulations and hence no longer any persistence (see Figure 2.2). For the case with a first convergence that would later spread for some reason (e.g. snow melt), the last convergence of the simulation period after which no spread occurred was used to estimate the persistence (Figure 2.2, experiment b)). The relaxation time [days] was the start of the simulation from experiment c) to the point of an equal oscillation of all simulations. All experiments a), b) and c) were repeated four times with a shift in the starting date from winter (January 1) to spring (April 1), summer (July 1) and fall (October 1). The starting date is the time where the initial conditions are set, i.e., the switch from warming up to prediction mode. All analyses were performed for each of the 100 parameter sets and for the persistence estimation as well as the catchment relaxation aggregated to a mean value in the end.

2.2.3 Importance of initial conditions vs. weather during prediction

The "prediction skill" of both experiment a) and experiment b) forecasts were calculated (Shukla and Lettenmaier, 2011). As reference, the reference simulation from experiment b) was used because it is the chronologically correct yearly sequence for each forecast/initial condition. Since we were interested in the effects on low flows, we based the measure of prediction skill of experiment a) (F_{ESP}) and experiment b) (F_{ESPrev}) on the absolute error as also used in F_{MARE} (eq. 2.2 and 2.3).

$$F_{ESP} = \frac{1}{n_{ic}} \sum |Q_{ref,b}(t,i) - Q_{sim,a}(t,i)| \tag{2.2}$$

where n_{ic} is the number of initial conditions (26 different years), $Q_{ref,b}(t,i)$ is the reference of the forecast i at day t and $Q_{sim,a}(t,i)$ is the ensemble member using the initial condition i at day t.

$$F_{ESPrev} = \frac{1}{n_{fc}} \sum |Q_{ref,b}(t,i) - Q_{sim,b}(t,i)| \tag{2.3}$$

where n_{fc} is the number of forcing ensemble members (26 different years) and $Q_{sim,b}(t,i)$ is the ensemble member at this day and forecast. The time dependent ratio F_{ratio} of F_{ESP}

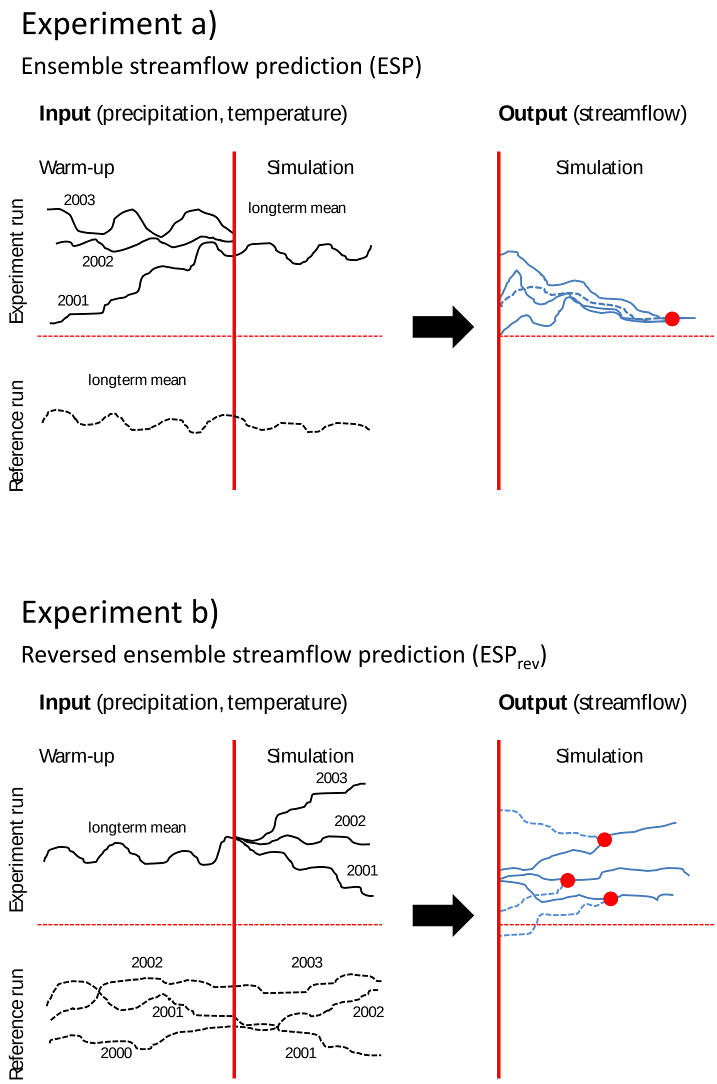

Figure 2.2: Set up of model experiment a) (ensemble streamflow prediction, ESP) and b) (reverse ensemble streamflow prediction, ESP_{rev}). Dashed lines indicate the reference runs and the red points indicate the persistence.

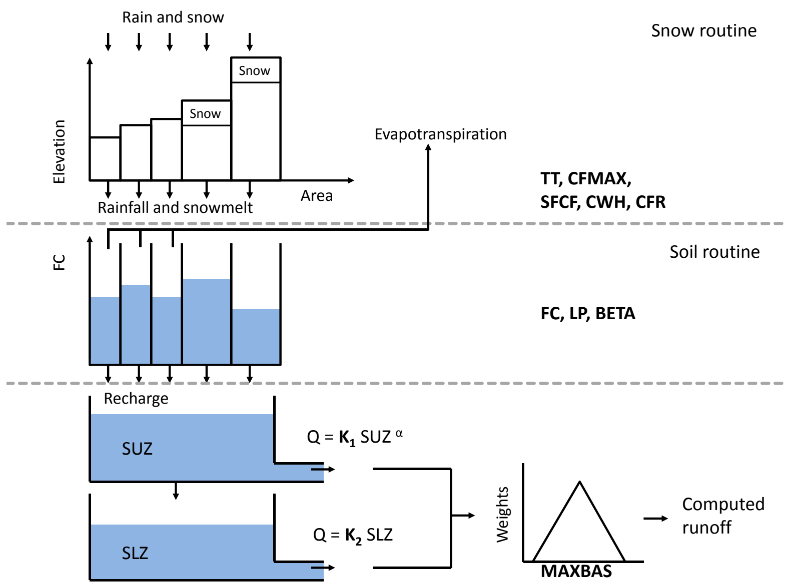

Figure 2.3: Conceptualization of the HBV model (modified after Uhlenbrook et al. (1999)).

and F_{ESPrev} of each experiment was then calculated using Equation 2.4.

$$F_{ratio}(t) = \frac{F_{ESP}}{F_{ESPrev}} \tag{2.4}$$

Values of F_{ratio} larger than one indicate a relatively higher forecast error due to uncertainties in the weather during prediction compared to the uncertainties in the initial conditions. This suggests a high contribution of the weather to the prediction skill (Shukla and Lettenmaier, 2011). Values of F_{ratio} smaller than one indicate relatively larger uncertainties due to the initial conditions compared to the uncertainties in the weather during predictions, which suggests a high contribution of the initial conditions to the prediction skill. The F_{ratio} of all simulations was calculated for lead times of 1, 2, 3, ..., 52 weeks. The values for F_{ratio} were computed for each of the 100 calibrated parameter sets and then aggregated as the mean.

2.2.4 Connection of persistence to conceptual storages

The HBV model consists of a number of conceptual storages: snow storage ($Snow$), soil moisture (SM), upper groundwater (SUZ), and lower groundwater storage (SLZ) (Figure 2.3). The initial storages at the start of each simulation were compared to the estimated persistences from experiment a). The actual initial hydrological state at the start of each simulation was transformed to a relative initial hydrological state by using the long term average conditions of the respective month in which the simulation start was set. For instance in winter the relative initial state is the ratio of the state on January 1 in a particular simulation and the average January state condition from the entire 26-year-period. The relation of initial conditions of each storage ($Snow$, SM, SUZ, and SLZ) from the 21 years of each catchment to the respective persistences were then analyzed by calculating the Spearman rank correlation between initial state and persistence for each catchment. Correlations with a p value smaller than 0.05 were considered statistically significant.

3 Results

3.1 General model performance

The model performance (F_{MARE}, eq. 2.1) of the best parameter sets varied between 0.64 and 0.84 for the 21 catchments with a median of 0.77. Good model performance could be achieved with varying individual parameter values and on average the best parameter values for a single catchment varied over 10 to 66 % of the tested parameter ranges.

3.2 Persistence in streamflow simulations

Experiment a) and b) resulted in similar estimates for persistence in streamflow for all catchments ranging between 50 days of persistence to more than a year (Figure 2.4). There was a tendency for higher elevation catchments to have longer persistences. We found strong correlations between the mean of the persistence estimates and the mean catchment elevations for all seasons (Table 2.2). The difference in estimates for the different starting dates was small. For spring and summer catchments 9 to 18 have higher persistence estimates for experiment a) than for experiment b). This difference is still visible for the values based on fall simulations, but is not apparent for the winter simulation. The variability of the persistence estimates caused by parameter uncertainty (i.e., the spread among the simulations of the 100 parameter sets) was higher than that caused by the inter-annual variability (i.e., the spread among the simulations for the different years) (Figure 2.5). Especially simulations starting in summer and fall showed an increased variability from parameter uncertainty for many catchments.

3.3 Catchment relaxation

The catchment relaxations varied between about three months to a year. For the low elevation catchments the catchment relaxation remained the same for all seasons, while the higher elevation catchments showed differences when starting the simulations at different dates. In Figure 2.6 the estimated mean persistences and the catchment relaxation times are compared. All catchments but catchment 18 have longer persistences than catchment relaxations. The difference between catchment relaxation and mean streamflow persistence was smallest in spring and became larger in summer, fall and winter. The largest difference between relaxation and persistence was seen in fall.

3.4 Importance of initial states vs. weather during prediction

F_{ratio} was found to vary with the season of the start of the simulation for the different catchments (Figure 2.7). For clarity, it should be mentioned again that the F_{ratio} indicates the relative influence of the initial conditions in comparison to the weather, while the persistence indicates the influence of the initial conditions on the predictions without specifying the role of the weather. In spring, the F_{ratio} with values smaller than one had

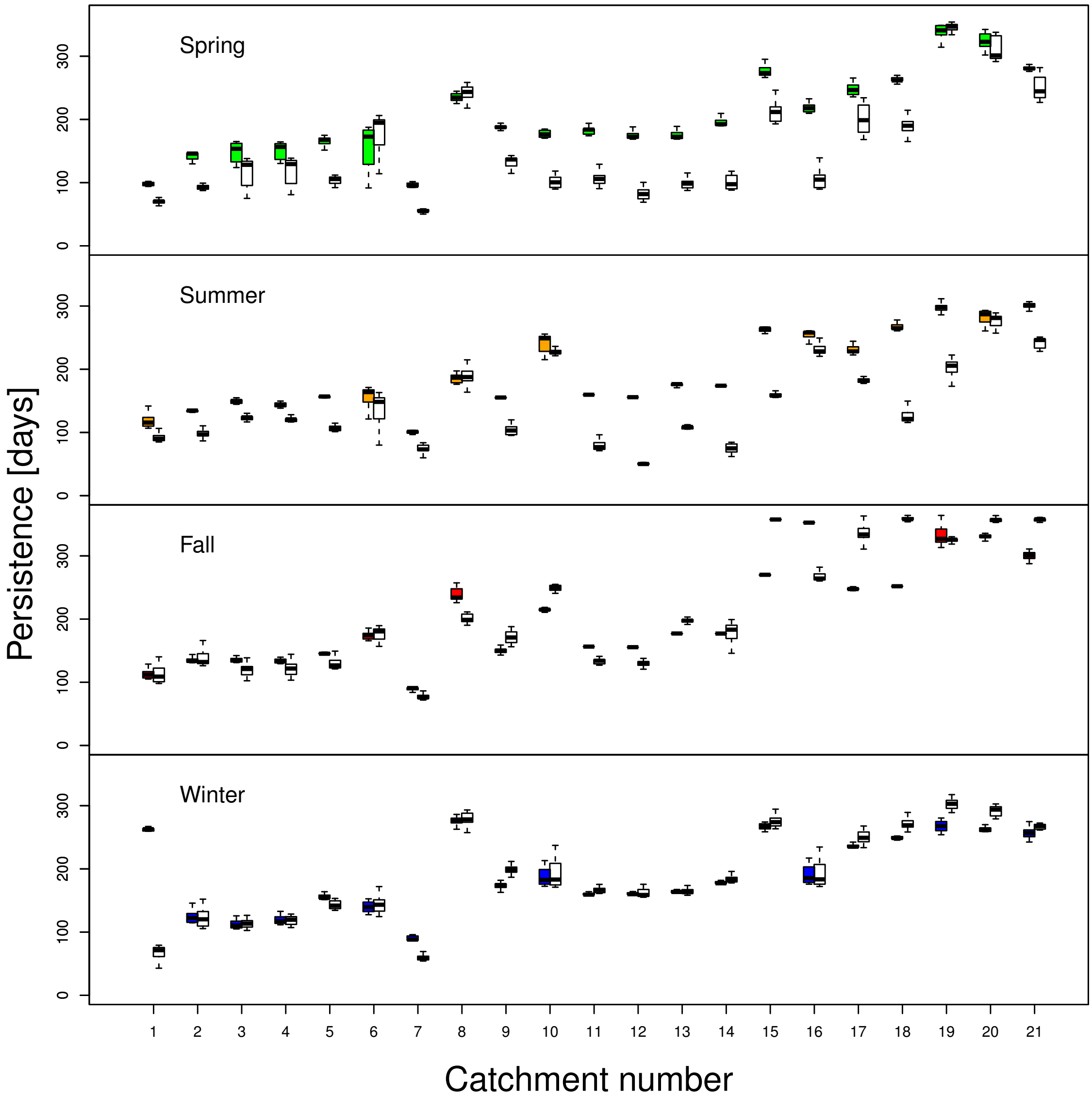

Figure 2.4: Distributions of the estimations of the persistences from experiment a) and experiment b) for the four starting dates for all catchments. For each catchment two distributions are displayed; the left colored box is the distribution from experiment a) and the right, empty box from experiment b). The catchment mean elevation increases from left to right.

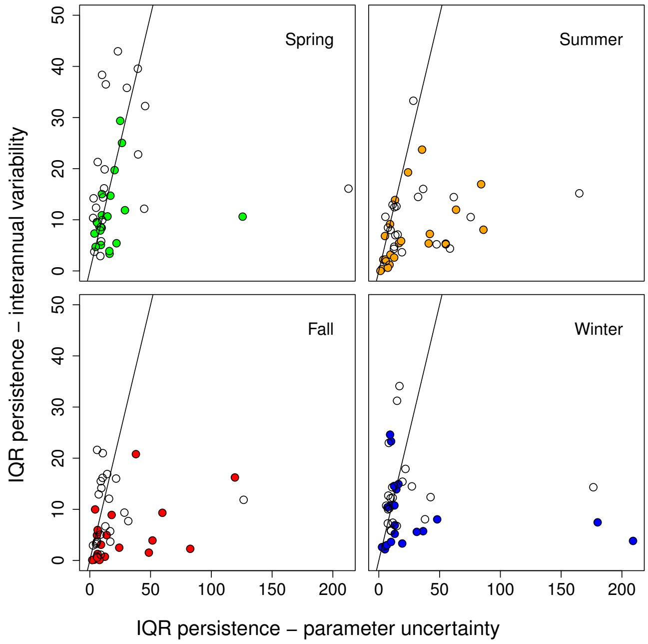

Figure 2.5: Comparison of the importance of variability of the estimated persistence values due to inter-annual variation and parameter uncertainty. The variability is quantified by the interquartile range, IQR, in the first case computed among the mean values from all 100 parameter sets and in the second case computed from the means of all 26 years. Colored symbols indicate the IQR resulting from experiment a), empty symbols from experiment b).

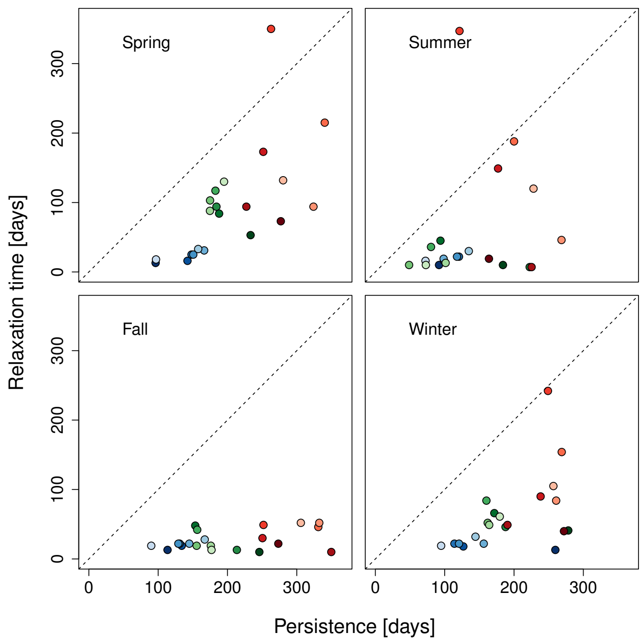

Figure 2.6: Catchment relaxation times compared to the mean persistence estimates (experiment a)). The colors range from blue for low elevation catchments to red for high elevation catchments.

Table 2.2: Spearman rank correlation between mean catchment elevation and mean of the persistence estimates from experiment a) and b).

Start of simulation	Spearman rank correlation	
	Experiment a	Experiment b
Spring	0.59**	0.90***
Summer	0.52*	0.90***
Fall	0.85***	0.89***
Winter	0.81***	0.60**

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

the longest lead times in most low elevation catchments and the highest elevation catchments with lead times ranging from 8 to 11 weeks. Middle and high elevation catchments have only very short lead times of about a week, during which the initial conditions have greater uncertainties than the weather during prediction. In summer, the length of the lead times with an F_{ratio} smaller than one varied for the catchments, but the pattern could not be clearly related to catchment characteristics. However, many low elevation catchments have an F_{ratio} smaller than one for lead times from 9 to 10 weeks. The shortest lead time with an F_{ratio} smaller than one when starting in summer was one week and the longest lead times with an F_{ratio} smaller than one 12 weeks. In fall, there is a clear tendency of greater uncertainties of the initial conditions for a longer period than those for weather for higher elevations. The shortest lead time with an F_{ratio} smaller than one when starting in fall was five weeks, and the longest lead time 18 weeks. In winter for all but the high elevation catchments the uncertainties of the initial conditions relative to the uncertainties of the weather decreased quickly and for most catchments with an F_{ratio} smaller than one, lead times were at the maximum one to three weeks. For the high elevation catchments the lead times with an F_{ratio} smaller than one ranged from 5 to 19 weeks, and for the three highest elevation catchments with an F_{ratio} smaller than one, the range was from 14 to 19 weeks.

3.5 Hydrological states and streamflow persistence

The main snow accumulation happens in early spring and winter. For most catchments, more snow during the initial conditions in winter were related to longer persistences (Figure 2.8). The Spearman correlation coefficients ranged between 0.46 and 0.66 for the statistically significant positive correlations in winter (Figure 2.10). In spring this relationship could only be found for a few catchments. Neither in spring nor in winter, could the catchments with significant correlations be attributed to the catchment properties. In summer only the highest elevation catchments would show snow effects, while in fall there might be single days of single years where snow starts to accumulate. For this reason we looked only at the relation between persistence and snow storage in winter and spring.

Drier initial soil moisture conditions in winter and spring for most catchments were related to longer persistences (Figure 2.9). The initial conditions of the other seasons showed both positive and negative correlations for different catchments (Figure 2.10). The negative correlations in spring and winter were found for low and middle elevation catchments. In summer the correlations could not be attributed to catchment properties.

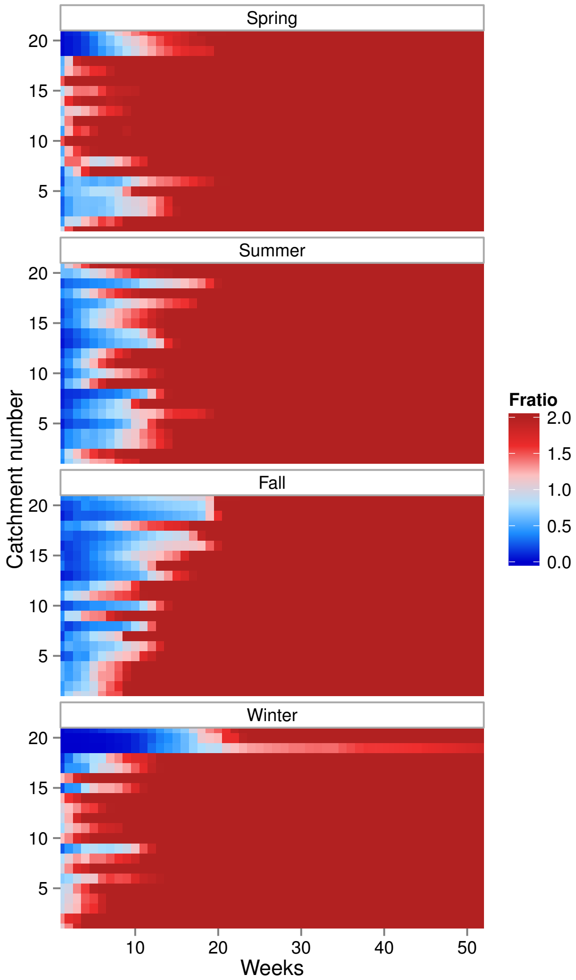

Figure 2.7: Median of F_{ratio} of the different simulations with lead times starting from one week up to one year in weekly time steps for all catchments and seasons. F_{ratio} values smaller than one, blue colors, indicate a larger uncertainty from the initial condition compared to the weather during the predictions; F_{ratio} values larger than one, red colors indicate a larger uncertainty stemming from the weather during the prediction.

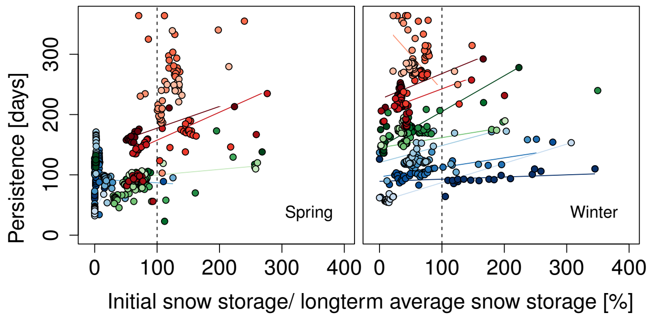

Figure 2.8: Relation of the initial snow accumulation relative to the snow accumulation during the simulation of the following year and the estimated persistence (experiment a)) for all catchments. Values below 100 indicate that the initial conditions were drier than the average snow accumulation during the simulation. Each color indicates a single catchment and each point a single year. The colors range from blue for low elevation catchments to red for high elevation catchments. For significant rank correlations linear regression lines are drawn.

However, the positive correlations in fall were mainly found for low elevation catchments, while the negative correlations were rather found for middle and high elevation catchments.

The initial conditions of the upper groundwater storage (SUZ) showed a clear tendency related to the persistence only in spring. Here, drier initial SUZ led to longer persistences for most catchments. There were significant correlations for the low and high elevation catchments, but not for the middle elevation catchments (Figure 2.10). The initial conditions in the other seasons were both positively and negatively correlated to the persistence. For winter only very few catchments showed significant correlations between initial conditions and persistence.

The lower groundwater storage (SLZ) with a simulation start in spring showed both significant positive and negative correlations to the persistence (Figure 2.10). In spring negative correlations were found for low elevation catchments, while the positive correlations did not match patterns of catchment elevation or size. In fall the positive correlations were found for the low elevation catchments, however, the negative correlations did not show any common pattern with catchment properties. In summer and winter, the correlations did not clearly match any catchment property pattern.

4 Discussion

4.1 Hydrological model

The results regarding persistence and relaxation times are to some degree model dependent. However, if a model has been successfully calibrated, differences are probably relatively small. It can be assumed that the important storages as well as their variability relative to each other are reasonably well represented. The model we used here was somewhat less complex than the VIC model (Liang et al., 1994), which has been used in several of the previous studies on persistence (Wood and Lettenmaier, 2008; Shukla and Lettenmaier, 2011). However, the groundwater routines of HBV and VIC are relatively similar. Using a less complex model allowed us to derive several behaviorable

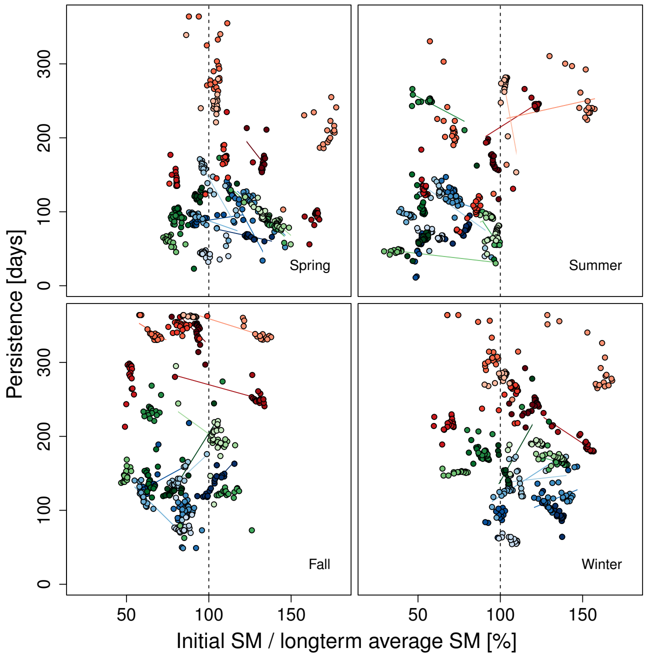

Figure 2.9: Relation of the initial soil moisture storage relative to the soil moisture storage during the simulation of the following year and the estimated persistence (experiment a)) for all catchments. Values below 100 indicate that the initial conditions were drier than the average soil moisture storage during the simulation. Each color indicates a single catchment and each point a single year. The colors range from blue for low elevation catchments to red for high elevation catchments. For significant rank correlations linear regression lines are drawn.

parameter sets and in this way to address parameter uncertainty, something that has not been done in the previous studies. From our results the use of an ensemble mean can be recommended, as the variability of the results due to parameter uncertainty was considerable for most of the catchments. The large variability among the simulations that were started in summer and fall when including parameter uncertainty indicates a high uncertainty connected to parameters of the soil routine which control evaporation. Seibert and McDonnell (2010) also concluded that it is important to consider parameter uncertainty to obtain reliable results. A high variability due to parameter uncertainty increases the risk for variable and partly random results if only a single parameter set is used. The ensemble approach used here is a suitable way to ensure robust results.

The simulated snow cover was derived from a degree day method, which could be argued to be less accurate than a snow cover simulated with energy balance methods. However, for the spatial and temporal scales looked at here, several studies have shown that the degree day method is an appropriate approximation (e.g., Rango and Martinec, 1995; Seibert, 1999; Hock, 2003; Merz and Blöschl, 2004)

The formulation of the potential evaporation can yield large differences in evaporative demand which can affect the calibrated model parameters and thus how the moisture is stored (McMahon et al., 2012). However, any errors in the estimation of the potential evaporation is implicitly considered in the calibration, i.e., parameter values might be influenced, but the catchment behavior in terms of responses and persistences should be

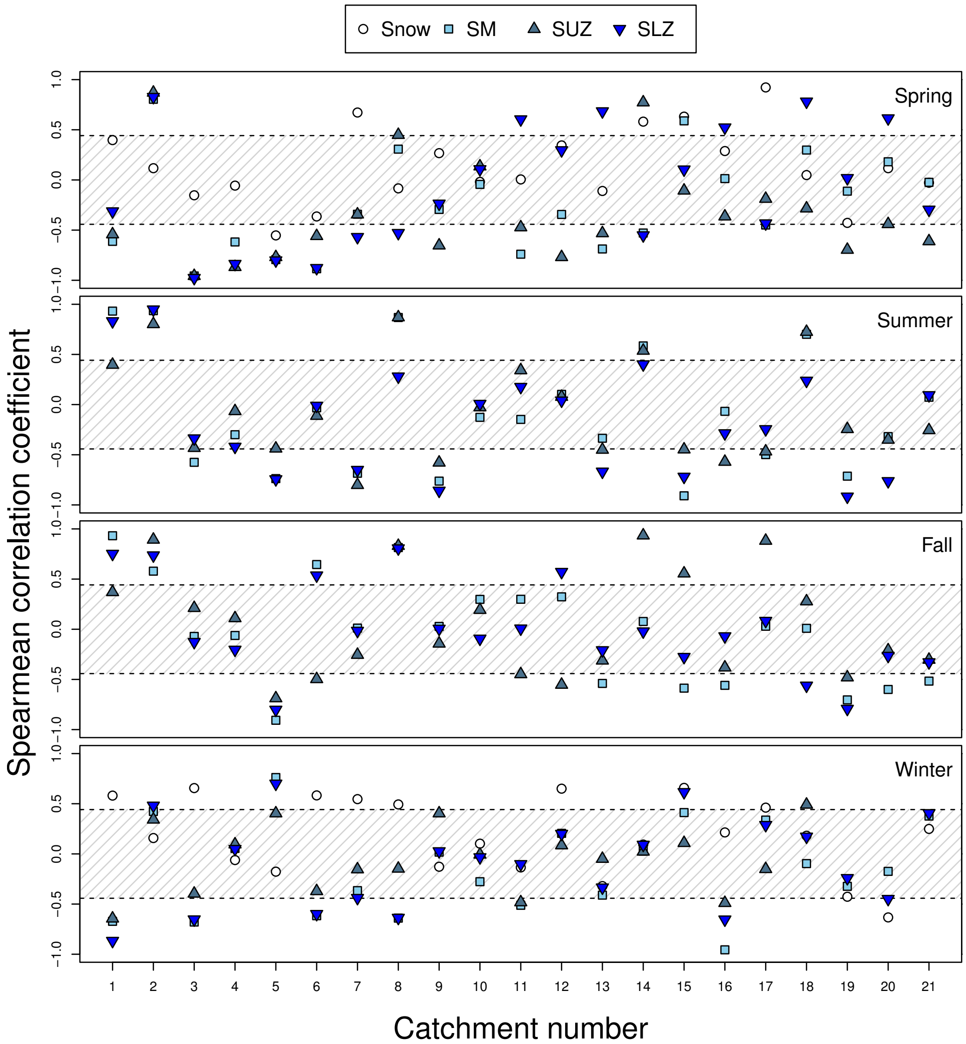

Figure 2.10: Spearman rank correlation coefficients for the relation between initial conditions of the storages (snow (*Snow*), soil moisture (*SM*), upper groundwater storage (*SUZ*) and lower groundwater storage (*SLZ*)) and persistences (experiment a)). Correlations that are not significant are plotted in the hatched area (p value >0.05).

influenced less. All these issues related to the model choice have to be considered, also when evaluating the results. However, the main outcomes concerning the influence of initial conditions related to storages within the catchment are represented and that the use of various parameter sets allowed for the estimation of uncertainty derived from the model.
Arithmetically averaged precipitation values were used in the climatology time series. While this approach ensures a representative mean precipitation amount, the temporal pattern of precipitation might be changed resulting in more days with some precipitation. During winter this has no effect on the simulated streamflow, but the mean simulated summer streamflow might decrease as more precipitation can be temporarily stored and then be evaporated. However, in the humid catchments used in this study, the effect on the total streamflow volume is limited. While it is important to be aware of this unrealistic temporal pattern in the precipitation climatology time series, its influence on the results of persistence and relaxation times in this study will not be substantial.

4.2 Prediction skill

Mahanama et al. (2012) started their simulations, as we did, in different seasons and looked at the ratios of the prediction skills (F_{ratio}) for several lead times up to six months. They found that depending on when the simulations were started and the lead time applied, the dominance of initial conditions or weather during prediction changed from more dominant initial conditions for short lead times (mostly 1 month) to more dominant weather during prediction for longer lead times. Mahanama et al. (2012) found that during spring and summer months initial conditions dominated the prediction skill in the U.S. beyond short lead times. We looked at the dominant effect at lead times up to one year and found at the shorter lead times relatively larger uncertainties stemming from the initial conditions and more uncertainties from the weather overall as compared to the initial conditions for all starting dates, which is similar to the findings of Mahanama et al. (2012) and the observations by Wood and Lettenmaier (2008) for the North Western US. However, the distribution of the F_{ratio} changed for different starting dates and for some catchments even more strongly. Shukla and Lettenmaier (2011) noted differences in the ratio of their objective functions for varying dry or wet initial conditions. We observed this as well, as the rather wet initial conditions in spring showed a dominant contribution of the weather during the prediction on the skill for the lower and highest elevation catchments. This changed for the drier initial conditions found in summer, where the uncertainties of the initial conditions are larger for longer compared to the uncertainties of the weather than in spring.

4.3 Variability of the persistence estimation

From the two experiments, the different model parameter sets for the simulations and the different seasonal forecasts as well as initial conditions, we found a distribution of persistence estimates for each catchment. The persistence estimates from experiment a) and experiment b) overlapped for most catchments. The persistence estimations from experiment b) were systematically longer in spring and summer for all catchments than the persistences from experiment a). In experiment a), the experiment run as such is a representation of what we face in reality, an attempt to forecast using a known initial condition and several scenarios of how the weather might be. By using reference simulations based on the true weather, this gave us the opportunity to see how long a present/initial state mattered in deriving the most realistic simulation rather than simply initializing the model with the climatology. Instead, in the reference of experiment a) both warming up and forcing was with the climatology. So, the persistences in experiment a) were computationally much faster to estimate than in experiment b) but the climatology plays a greater role in the definition of the persistence estimation. The role of climatology could be the reason for the observed offset in the persistence estimates for the middle elevation catchments: If the initial conditions were wetter and/or more snow accumulation took place during the models warming up phase, it would take longer to reach the reference simulation that was based on a drier climatology than it would take to reach a reference simulation that was based on a realistic seasonal warm up (as in the reference runs from experiment b)). For fall and winter simulations the climatology was likely closer to a realistic seasonal warm up, since we could not observe this offset for those seasons.

4.4 Streamflow persistence vs. catchment relaxation

The estimated streamflow persistences are a combination of both weather and catchment properties. Catchment relaxation times should instead mainly represent the catchment storage properties. The relaxation times in different seasons however can vary slightly as the simulations started with different initial conditions each season and then reached a new balance of the system. The catchment relaxation time for catchments with a snow dominated streamflow regime were longer in spring compared to the other seasons, which could be explained by filled soil and groundwater storages from the preceding winter and fall. Since the lower elevation catchments did not show this seasonal difference we suspect the higher catchments to have larger storages.

4.5 Initial conditions and catchment properties

We found that the persistence estimates were strongly correlated to catchment mean elevation. This could partly be explained by an increasing snow influence with elevation, but could also be due to larger aquifers. In the synthetic experiments of Van Loon et al. (2014), who compared warmer and colder climates as well as the effect of varying geology, both increased snow influence and slower aquifer response were found to cause longer drought persistences. In our study, we also found that initial storages of snow and soil moisture were related to the persistence estimates, which corresponds to the conclusion of Van Loon et al. (2014) that seasonality effects cannot be explained by meteorological processes alone. The relation between storage of snow/soil moisture and persistence was also found by Singla et al. (2012) for France and by Mahanama et al. (2012) for the U.S.. While Singla et al. (2012) could distinguish between the importance of snow and soil moisture for elevation classes, we did not find such a clear signal. Instead we saw that the importance of snow, soil moisture and groundwater storage, depended on the starting date of the simulations. When the simulations were started in winter or spring the initial conditions of snow were related to the persistence estimates for many catchments and in summer to the highest with more initial snow leading to longer persistences. Drier initial soil moisture could be connected to longer persistences for lower elevation catchments with simulation starts in all seasons but winter. Longer predictabilities connected to drier initial conditions were also found by Fundel et al. (2013). For higher elevation catchments and winter simulation start wetter initial conditions lead to longer persistences. This can be explained by the absolute size of the soil moisture storage of lower elevation catchments compared to higher elevation catchments. The persistences and initial groundwater storage conditions did not show a general pattern.

4.6 The role of snow

Accumulating and melting snow is an important storage and storage outflow. Moreover, snow melt fills other storages in the catchment. Hence, when trying to distinguish between meteorological influence and initial conditions with the ESP/ESP_{rev} analysis this double role of snow has to be taken into account. Snow melt that contributed to the initial conditions is attributable to the initial conditions, but snow fall, accumulation and melt during the simulation period will directly influence the meteorological forcing. The high elevation catchments where snow fall could also occur in seasons other than winter showed a different effect than the catchments at middle elevations, where the initial conditions were still more dominant than the meteorological forcing. This could result from the time shift of when the snow accumulation and melt happened.
For the persistence estimation, snow storage is directly taken into account, which was visible in both the correlation to the mean catchment elevation and the relation between snow storage and persistence. For the catchment relaxation, the direct snow accumulation and melt was explicitly excluded, even though the snow melt that occurred during the warm-up was included. This remaining snow influence seems critical as we found seasonal differences in the relaxation times of the middle and higher elevation catchments, but not in the lower elevation catchments.
Another indication for the role of snow can be seen from the already discussed offset between the results from experiment a) and b), namely that the climatology in the warming up of the reference runs in experiment a) were not as realistic as the warm up of experiment b), which caused greater offsets in the seasons with snow involved.

4.7 Catchment elevation and storage

At high elevations we usually find thinner soils, however, our results challenge the common assumption of less storage in higher elevation catchments and indicate that there might be a larger groundwater storage. This can be explained by large storage features that can be found in mountain catchments like talus slopes with high storage capacities. The total storage capacity might also increase with elevation because of a storage volume above drainage level that is higher in mountainous catchments than in rather flat low elevation catchments. We know for example that the highest catchment (catchment 21) from our selection shows extraordinarily high storage capacities as water can be stored in deep moraines that make up one third of the entire area and in an additional alluvial storage on the valley floor (FOEN, 2011).

4.8 Predictability of droughts

In this study, the analyses were performed from a low flow perspective, as the objective function during both the calibration and the analysis emphasized low flow. The persistence estimations showed that for different catchments the maximum predictability for streamflow varied from 50 days to more than a year with the tendency to show higher elevation catchments related to longer predictabilities. The persistence estimates did not vary greatly with a change of the starting date of the simulations to another season. The relative influence from weather with respect to initial conditions, however, varied with a change of the starting date of the simulations. In spring the highest elevation catchments had longer lead times with small uncertainties of the initial conditions presumably due to large snow accumulations at the start of the simulations for all years of the ensemble. The lower elevation catchments, however, have, at the time of the start of the simulation, barely accumulated snow, while the snow storage at middle elevation catchments might vary strongly from year to year. This can explain the longer relative influence of the initial conditions on the predictability found for the low elevation catchments, but not in the middle elevation catchments, as the snow can accumulate before or after the starting date of the simulation (April 1). In fall and winter higher elevation catchments tended to have longer lead times of high relative importance of the initial conditions compared to the weather during prediction. This points to a larger influence of the initial conditions in higher elevations which could be due to snow storage as well as other storages. With the tendentially drier conditions in summer there was more variation and the simulations of catchments, no matter at which elevation, had longer or shorter small uncertainty contributions from the initial conditions. The summer F_{ratio} point on the one hand to storage differences, but also to varying summer meteorology for the different catchments. With this study, the question of how long before a drought occurs can it be predicted, cannot readily be answered. However, for the catchments in this study we found ranges of maximum detectable influence of initial conditions from 50 days to more than a year. Further, we found that the catchment elevation matters more than the starting date of the simulation for a maximum predictability of streamflow and that the relative importance of initial conditions compared to the relative influence of the weather during the predictions changes with the season in which the simulation start is set.

5 Conclusions

We estimated persistences for 21 different Swiss catchments using model simulation experiments performed with the HBV model. The range of the persistence estimates differed between the catchments and showed a strong correlation with mean catchment elevation. Together with the relative influence of weather with respect to initial conditions, the predictabilities ranged from 50 days to more than a year with a decreasing influence of the initial conditions over time. The degree of the decrease was found to be dependent on the start of the simulation. In fall and winter, a longer influence of the initial conditions during prediction was found for higher elevation catchments as compared to the weather. In spring, the initial conditions were relatively more important for the prediction than weather for the highest and lower elevation catchments compared to

the middle elevation catchments. This might be due mainly to annual snow melt and accumulation variations around the starting date of the spring simulations in the middle elevation catchments. In summer, the initial conditions had differing influence on the predictions and were not related to a specific elevation range.
The interpretation of the correlation between higher elevation and longer persistences might not be easy without additional information about catchment properties like type and size of aquifers. Compared to the persistence the relaxation time was lower and the catchment relaxation time varied seasonally for higher elevation catchments but was constant for lower elevation catchments, which indicates the important role of snow in persistence estimation. We found that snow and soil moisture as well as groundwater initial conditions derived from the model states were related to the persistence estimates. Drier initial states of soil moisture and groundwater and more snow accumulation at the start of the simulation led to longer persistence estimates.
In opposition to an intuitive expectation from shallow soils in higher elevations, we found an indication for larger groundwater storages in higher elevation catchments. This may motivate a reconsideration of the sensitivity of mountainous catchments to low flows in a changing climate.

APPENDIX 3

A drought index accounting for snow*

Abstract

The Standardized Precipitation Index (SPI) is the most widely used index to characterize droughts that are related to precipitation deficiencies. However, the SPI does not always deliver the relevant information for hydrological drought management particularly in snow influenced catchments. If precipitation is temporarily stored as snow, then there is a significant difference between meteorological and hydrological drought because the delayed release of melt water to the stream. We introduce an extension to the SPI, the Standardized Snow Melt and Rain Index (SMRI), that accounts for rain and snow melt deficits, which effectively influence streamflow. The SMRI can be derived without snow data, using temperature and precipitation to model snow. The value of the new index is illustrated for seven Swiss catchments with different degrees of snow influence. In particular for catchments with a larger component of snowmelt in runoff generation, the SMRI was found to be a worthwhile complementary index to the SPI to characterize streamflow droughts.

1 Introduction

Droughts always originate from a lack of precipitation. In some regions high temperatures and evapotranspiration are additional important drivers of soil moisture and hydrological droughts. In contrast to these drought processes that occur in summer, the storage of precipitation as ice and snow can act as a key moderator of hydrological drought. In particular, streamflow droughts are often related to the presence or absence of snow in the preceding winter period and winter droughts can occur despite large amounts of precipitation, if the precipitation falls as snow. Van Loon and Van Lanen (2012) distinguish between six different hydrological drought types according to their development (classical rainfall deficit drought, rain-to-snow-season drought, wet-to-dry-season drought,

*Accepted for publication in Water Resources Research as: Staudinger, M., Stahl, K., and Seibert, J., A drought index accounting for snow.

cold snow season drought, warm snow season drought, and composite drought). Since hydrological droughts can have severe impacts on river ecology, water supply, energy production, or navigation, there is a need to monitor these droughts.

Drought monitoring requires indicators that are general enough to be widely applicable, but specific enough to capture the type of drought relevant to the region and variable of interest. The development of such indicators in the United States is summarized by Heim Jr (2002). There are only a few indices that consider snow explicitly, one of these for example, is the surface water supply index (SWSI) (Shafer and Dezman, 1982; Doesken et al., 1991). The Standardized Precipitation Index (SPI) is an indicator for drought that was first introduced by McKee et al. (1993). Since its introduction, the SPI has been applied in many studies, in operational drought monitoring in the present, and also in scenario predictions of drought for climate change impact assessment (e.g. Ji and Peters, 2003; Ghosh and Mujumdar, 2007; Naresh Kumar et al., 2009; Orlowsky and Seneviratne, 2012; Naresh Kumar et al., 2009). A major advantage of the SPI compared to other drought indices is that it requires only precipitation data to describe drought severity. It is calculated based on a theoretical probability distribution fitted to the long-term precipitation record aggregated over a chosen preceding period. This probability distribution is then transformed into a normal distribution so that the mean SPI is zero. Positive SPI values indicate greater than mean precipitation, and negative values indicate less than mean precipitation. As the SPI is standardized, wetter and drier climates are represented in the same way allowing for regional comparison studies (Hayes et al., 1999).

Different precipitation aggregation periods can reflect the impact of drought as it propagates through the hydrological cycle into soil, streamflow and groundwater. Soil moisture conditions are related to precipitation anomalies on a relatively short scale, whereas streamflow for instance, reflects longer-term precipitation anomalies (Hayes et al., 1999). With the right aggregation time a climatic drought index such as the SPI may also be a suitable indicator for hydrological droughts. The US Drought Monitor, for example, uses composite drought indices with a focus on short SPI aggregation periods for warnings about agricultural drought impacts and composite indices with a focus on longer SPI aggregation periods for warnings about hydrological drought impacts (droughtmonitor.unl.edu). Several studies have investigated the time lag between SPI and streamflow drought in order to find the most suitable SPI aggregation period linked to hydrological drought characterization. Some researchers have determined such a time lag between meteorological drought and streamflow drought (Haslinger et al., 2014), while others found strong dependencies apart of areas that have a large groundwater storage (Haslinger et al., 2014) or at times of snow storage (Shukla and Wood, 2008; Vidal et al., 2010).

To create a methodologically consistent indicator of hydrological droughts, several studies have transferred the SPI approach to observed and modeled hydrological variables. López-Moreno et al. (2009) and Vicente-Serrano et al. (2011) applied the SPI concept to observed streamflow in Spain, introducing a standardized streamflow index (SSI). Shukla and Wood (2008) derived a standardized runoff index (SRI) for monthly aggregations of modeled daily grid cell runoff, which consisted of modeled surface runoff and base flow (subsurface flow). The results were SRI maps for the entire USA based on the grid cells of a large-scale hydrological model. Vidal et al. (2010) applied the approach to hydrological model output for France, but instead of grid cell runoff they calculated a

Table 3.1: Catchment properties (FOEN, Section Hydrology, 2011).

Catchment number	Name	Area (km^2)	Mean elevation (m a.s.l.)	Elevation range (m a.s.l.)	Regime	Snow/precip[a] (%)
1	Mentue	105.0	679	445-927	pluvial	4.8
2	Sense	352.0	1068	548-2189	pluvio-nival	11.7
3	Sitter	74.2	1252	769-2501	nival	22.7
4	Allenbach	28.8	1856	1297-2762	nival	33.7
5	Riale di Calneggia	24.0	1996	885-2921	nival	34.3
6	Ova da Cluozza	26.9	2368	1508-3165	nival	42.2
7	Dischma	43.3	2372	1668-3146	nival	44.7

[a] Percent of snow in precipitation is calculated as the ratio of precipitation on days with air temperatures below 0°C and precipitation from the entire observational period

standardized flow index for the routed streamflow. Shukla and Wood (2008) and Vidal et al. (2010) compared their derived hydrological indices with the traditional SPI in order to explore the time lag of the drought propagation through the hydrological cycle. They concluded that a standardized runoff index can complement the SPI especially in periods when variables other than precipitation become more important, e.g. periods of snow accumulation and melt. While the advantage is that modeled runoff considers precipitation, temperature and radiation as well as information about the variability of vegetation, soil and terrain characteristics, it cannot be validated. Only runoff routed to the outlet of a catchment, i.e. the streamflow, can be gauged and thus validated. Unfortunately, in many catchments, streamflow data are influenced by human impacts or are not available for periods long enough to calculate an SSI based on observations.
The SPI can be modified to indicate a hydrological drought rather than a precipitation drought without the full complexity of a hydrological model, by only accounting for first-order controls on catchment hydrology that affect drought. Recently, Vicente-Serrano et al. (2010) introduced an index that accounts for evapotranspiration as an important amplifier of drought and found this index to be useful for catchments in Spain. This study specifically aimed for a climatic drought index with low data requirements which can serve as an indicator for hydrological drought in regions with a variable influence of snow. In such regions, e.g. mountain headwaters, streamflow is a major source of water use for water supply, energy production, and the ecology of mountain streams is vulnerable to drought. Therefore, this study uses a drought index based on observed streamflow, the SSI, as a benchmark against which to compare the climatic drought index SPI and the new Standardized Snow Melt and Rain Index (SMRI). The comparison is done for seven Swiss catchments with different amounts of snow melt contributions to streamflow.

2 Data and Methods

2.1 Data

Data from seven unregulated meso-scale catchments in Switzerland were used in this study (Table 3.1). The mean elevation for the different catchments ranges between 700 and 2400 m a.s.l. The catchment areas range between 20 and 350 km^2, and the estimated fraction of annual snow in precipitation ranges between 5 and 45% (Table 3.1). Daily precipitation and temperature data were derived from the grid products RhiresD and TabsD (Frei, 2013) provided by MeteoSwiss (2013). Both grid products are based on the interpolation of the daily anomalies of a dense network of meteorological records on a spatial background climatology. The daily grids have a spatial resolution of 2km x 2km and cover the period 1971-2011. For this study, catchment averages of precipitation and temperature were computed. Observed time series of daily mean streamflow were available for the same period (1971-2011) for all catchments. (FOEN, 2012).

2.2 Probability distribution selection for SPI and SSI

For the calculation of standardized drought indices a theoretical distribution has to be chosen. The SPI has often been calculated based on the Gamma distribution, even though some authors claim that other distributions like the Pearson type III distribution might be more suitable (e.g., Guttman, 1999). We tested different theoretical distributions as suggested by Vicente-Serrano et al. (2010) and Vicente-Serrano et al. (2011). The best fit for all variables on average was found for the Pearson type III distribution, which then served as a basis for all index calculations (SPI, SSI and the new SMRI). The parameters of the distribution were estimated using the L-moments method as described by Hosking (1990).

2.3 The Standardized Melt and Rainfall Index (SMRI)

The new SMRI was calculated similarly to SPI and SSI, but from the daily sum of snow melt and rain (MR). To obtain daily snow melt amounts, a commonly used snow model that only requires temperature data in addition to precipitation was first applied. While any snow model or derivation of snow melt could be used to calculate the index, the model used here consists of a snow accumulation component based on a threshold temperature and of a snow melt component based on a degree-day approach allowing for storage of up to 10% of the current simulated snow water equivalent and refreezing of liquid water in the snow pack (at a reduced rate compared to melting) (e.g., Bergström et al., 1992) (a detailed description can be found in Appendix 5). The variable MR was then transformed into the index SMRI using the Pearson type III distribution.
In order to explore the level of local parameterization needed, three parameter set ensembles were tested: the first parameter set ensemble (Set 1) assumed no prior knowledge (10'000 random parameter sets), the second set (Set 2) assumed some regional knowledge and the third set (Set 3) assumed specific catchment knowledge. Set 1 was derived from a Monte Carlo analysis, where 10'000 parameter combinations were tested for the snow model. The sample for the Monte Carlo simulations was created using Sobol' sequences (R Package randtoolbox, CRAN, 2012). For the parameter sets we chose typical param-

eter ranges (Seibert, 1999) for the threshold temperature between -2.5 and 2.5 $°C$, for the degree-day factor between 1 and 6 $mm°C^{-1}day^{-1}$ (Esko, 1980; Seibert, 1999; Hock, 2003; Merz and Blöschl, 2004), for the refreezing coefficient between 0 and 0.1.
Set 2 and 3 came from calibrating a full hydrological model, which contains apart of the same snow model routine also soil and groundwater response and routing routines (HBV model in the version HBVlight (Seibert and Vis, 2012)). The model was automatically calibrated to observed streamflow for each catchment over the period 1971 to 2011. For the calibration a genetic optimization algorithm with subsequent steepest gradient tuning (Seibert, 2000) was used. 100 calibration trials were performed, which resulted in 100 optimized parameter sets for each catchment according to a combination of Nash-Sutcliffe model efficiency and volume error (Lindström et al., 1997b), where the weighting factor for the volume error was set to 0.1, as recommended by Lindström et al. (1997b) and Lindström (1997). The same parameter ranges that were used in the Monte Carlo simulations for Set 1 were used for the calibration. Set 2 was the resulting 100 optimized parameter sets for each catchment. Finally, for each catchment a so called regional parameter set (Set 3) was composed, consisting of Set 2 of all other catchments (i.e., here 100 x 6 =600).
These snow model parameter values were then used to compute MR and subsequently SMRI.

2.4 Application and comparison of SMRI to SPI and SSI

All indices were calculated for different aggregation periods (1, 2, 3, 4, 6 and 12 months), referred to as, for instance SMRI-6 for the SMRI calculated based on a six months preceding aggregation period. If no aggregation period is specified results refer to all aggregation periods.
To compare the new SMRI as well as the SPI to our variable of interest, the SSI, a benchmark model efficiency F_{bench} (Eq. 3.1, (Schaefli and Gupta, 2007)) was used as one measure of comparison. F_{bench} was calculated as the ratio of the quadratic absolute errors; subtracting the ratio from one transforms it to a range of minus infinity to one. A value of one for F_{bench} corresponds to a perfect fit of the SSI and SMRI. Values larger than zero indicate that the SMRI is closer to the SSI than the SPI and values below zero indicate that the SPI is closer to the SSI than the SMRI.

$$F_{bench} = 1 - \frac{\sum (x_{SSI}(t) - x_{SMRI}(t))^2}{\sum (x_{SSI}(t) - x_{SPI}(t))^2} \tag{3.1}$$

F_{bench} was calculated for both the entire index time series (1971-2011) as well as for the hydrological dry periods only (SSI <0).
In addition to this general evaluation, we looked at two historical drought events in particular: the summer drought of 2003 (Rebetez et al., 2006) as well as the spring drought of 2011. The summer drought 2003 was caused by a lack of precipitation and, due to extremely high temperatures, also high evapotranspiration rates. The drought in spring 2011, resulted from a preceding winter with little precipitation and thus little snow accumulation in combination with relatively high temperatures during spring time.

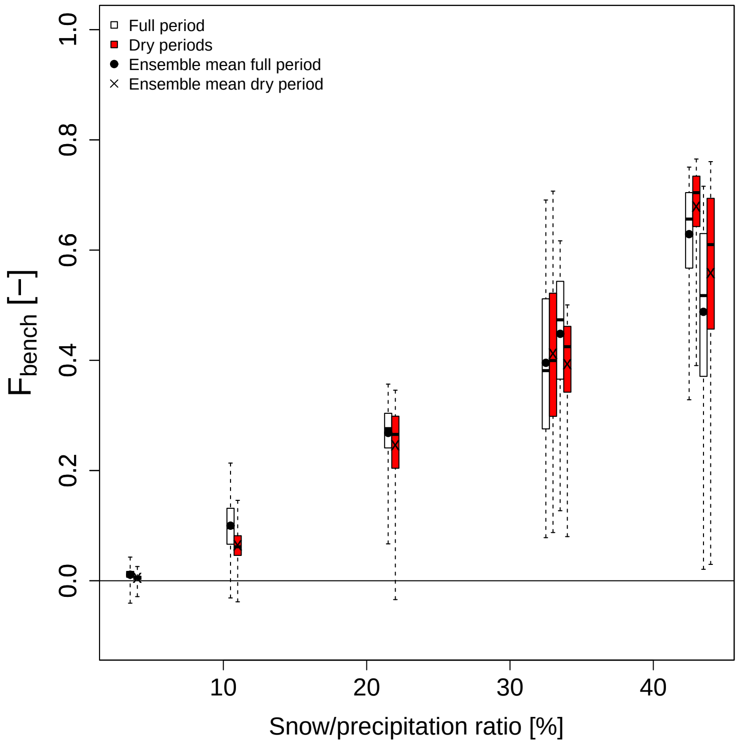

Figure 3.1: Distributions of the measure of comparison F_{bench} for snow model parameter Set 1 for drought indices with an aggregation period of three months. Each pair of boxes (white plus red) represents one catchment. Additionally, the measure of comparison of the ensemble mean is shown. The whiskers of a boxplot extend to the minimum and the maximum values, the box extends from the 25th to the 75th percentile and the bar shows the median.

2.5 Sensitivity to elevation distribution

The SMRI series were first computed for the mean catchment elevations. To assess how representative this lumped approach is, the SMRI computation was repeated in a semi-distributed way: each catchment was divided into elevation zones of 100 m. For each elevation zone both the fraction of the elevation zone of the catchment as well as the temperature change according to a fixed lapse rate of 0.6 $°C/100m$ were calculated. From the area-weighted mean of MR the SMRI$_{elev}$ was derived. Finally, the SMRI$_{elev}$ was compared to the SMRI for aggregation periods of one and three months using F_{bench} (Eq. 3.1). While the consideration of elevation zones changes the temporal distribution of snow accumulation and melt, for aggregation periods that are longer than the annual snow period this has no significant impact.

3 Results

The values for F_{bench}, derived from Set 1, were in most catchments and for most parameter sets greater than zero, which means that the SMRI was closer to the SSI than the

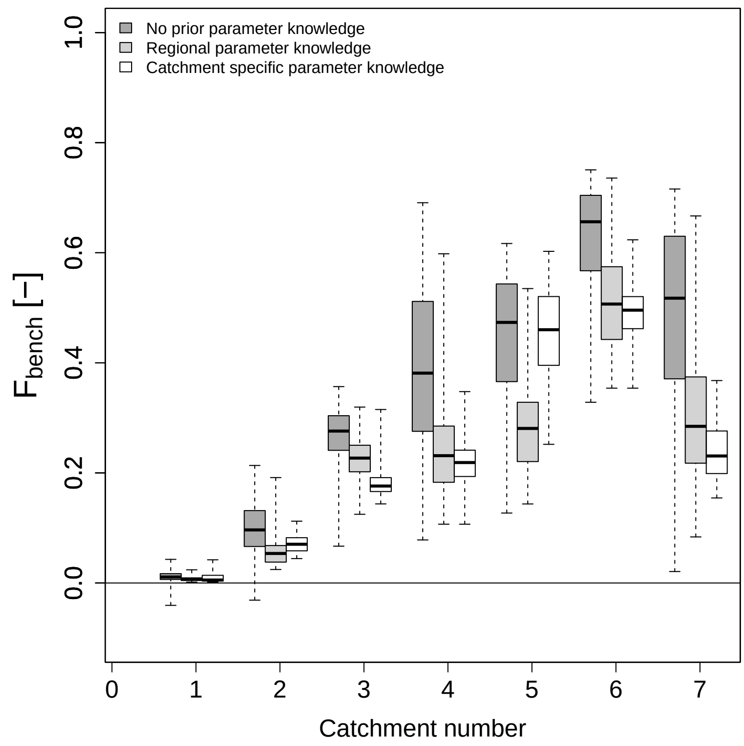

Figure 3.2: Distributions of the measure of comparison F_{bench} for the three different snow model parameter sets for drought indices with an aggregation period of three months. Boxplots as in Figure 3.1.

SPI for both the entire period and the dry periods (Figure 3.1). For the catchment with the smallest snow/precipitation ratio, the SPI and the SMRI were comparable. However, the difference increased systematically with increasing snow influence on the streamflow regime for both the entire period as well as for the dry periods only (Figure 3.1). The values for F_{bench} were on average slightly lower for the simulations which were based on parameters with prior knowledge (Sets 2 and 3), and the spread was smaller (Figure 3.2).

There were also seasonal patterns in F_{bench} (Figure 3.3): for the two catchments with the highest average snow contribution (>30%), F_{bench} decreases slightly in the summer months. For the catchments with between 10% and 23% average snow ratio, during the melt period (April, May and June) the hydrological droughts were closer represented by the SMRI than by the SPI. The Mentue catchment with a pluvial streamflow regime shows a closer representation of the SSI by the SPI in January and February while for the rest of the year by the SMRI.

Figures 3.4 and 3.5 show the droughts of 2003 and 2011 for the Ova da Cluozza catchment (nival). In 2003, the SMRI was closer to the SSI than the SPI regardless of the aggregation period. However, the ensemble mean overestimated the streamflow drought for the aggregation periods of one to four months. SMRI-6 and SSI-6 were similar regarding both severity and duration. The duration of the drought was captured well for all aggregation periods. While the SPI indicated severe droughts with values below -1.5,

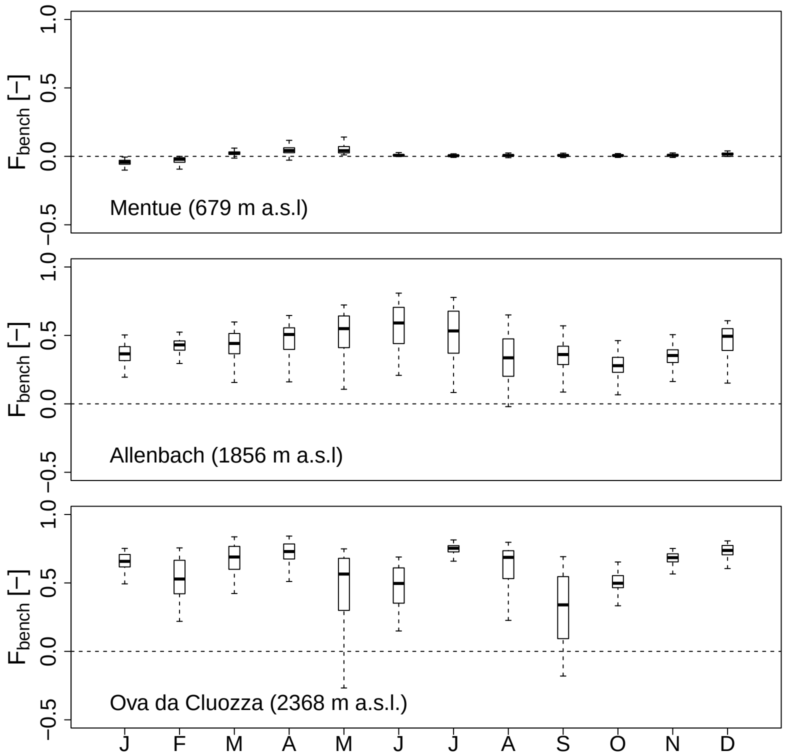

Figure 3.3: Distributions of the measure of comparison F_{bench} for each month, modeled with Set 1 for a catchment with little snow influence (upper), a catchment with medium snow influence (middle) and a catchment with high snow influence (lower). Boxplots as in Figure 3.1.

both SSI and SMRI indicated a less severe drought. For 2011, the ensemble mean of the SMRI mimicked the SSI in all aggregation periods. Here again, the SPI indicates more severe droughts than the SSI and SMRI.

Figures 3.6 and 3.7 show onset and end of the droughts in all catchments. While in the Mentue and Sense catchments SPI-3 and SMRI-3 fail to identify the start and end of the hydrological summer drought of 2003 as indicated by the SSI-3, for the nival Sitter, Allenbach and Riale di Calneggia catchments they better describe the start and end of the drought. For the two catchments with the highest elevation (Ova da Cluozza and Dischma) the SMRI matches the end of the drought as indicated by the SSI-3, but defines its start later. However, the SMRI-3 indicates the start of the drought about 1 month earlier, i.e. closer to the SSI than the SPI-3.
For the Mentue and Sense catchments, which generally have little snow, neither the SPI nor the SMRI capture the timing of the hydrological spring drought in 2011; also for the Sitter and Allenbach catchments, SPI and SMRI are similar. For catchments with the most snow the SMRI closer matches SSI than the SPI regarding the start and the end of the spring drought. Different thresholds that define different severities of droughts could be applied, which also bring the SMRI closer to the SSI than the SPI.
Including different elevation zones of a catchment improved SMRI-1 and SMRI-3 (Fig.

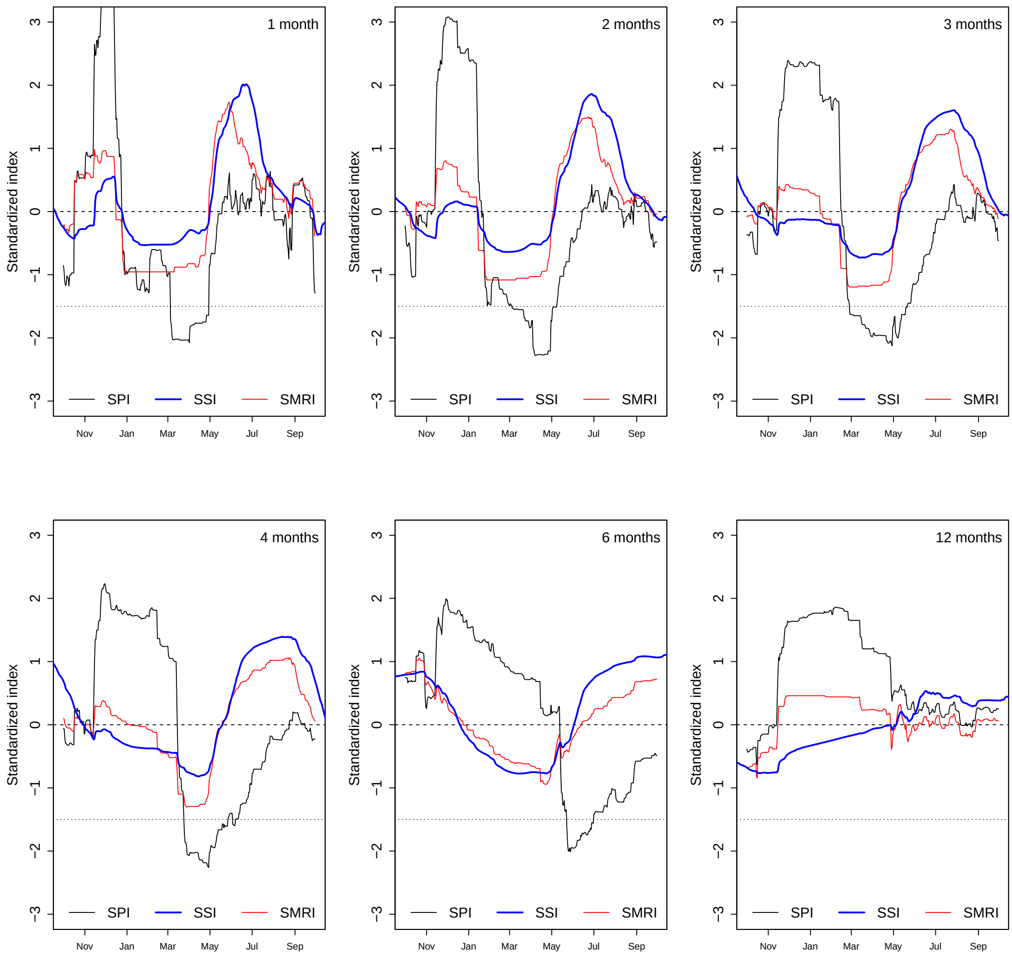

Figure 3.4: Standardized precipitation (SPI) (black), streamflow (SSI) (blue) and ensemble mean of the snow melt rain index (SMRI) (red) in daily resolution for six different accumulation periods during the summer drought 2003 for the nival Ova da Cluozza catchment.

3.8). The strongest improvement was found for catchments with mean catchment elevations from 1000 to 2000 m a.s.l. (Sense, Sitter, Allenach, Riale di Calneggia). However, the relative ranking of the catchments' F_{bench} is similar for SMRI and SMRI$_{elev}$. For the SMRI-1 the improvement when using elevation zones was slightly higher than for the SMRI-3. For both SMRI-1 and SMRI-3 a clear reduction of the spread in values was found when different elevation zones were considered.

4 Discussion

4.1 Uncertainties from model and index standardization

The proposed SMRI is an index that is calculated in a two-step process; i.e. first a model is applied that accounts for the dominant process that affects severity and timing of hydrological drought, and then the output of this model is transformed into the in-

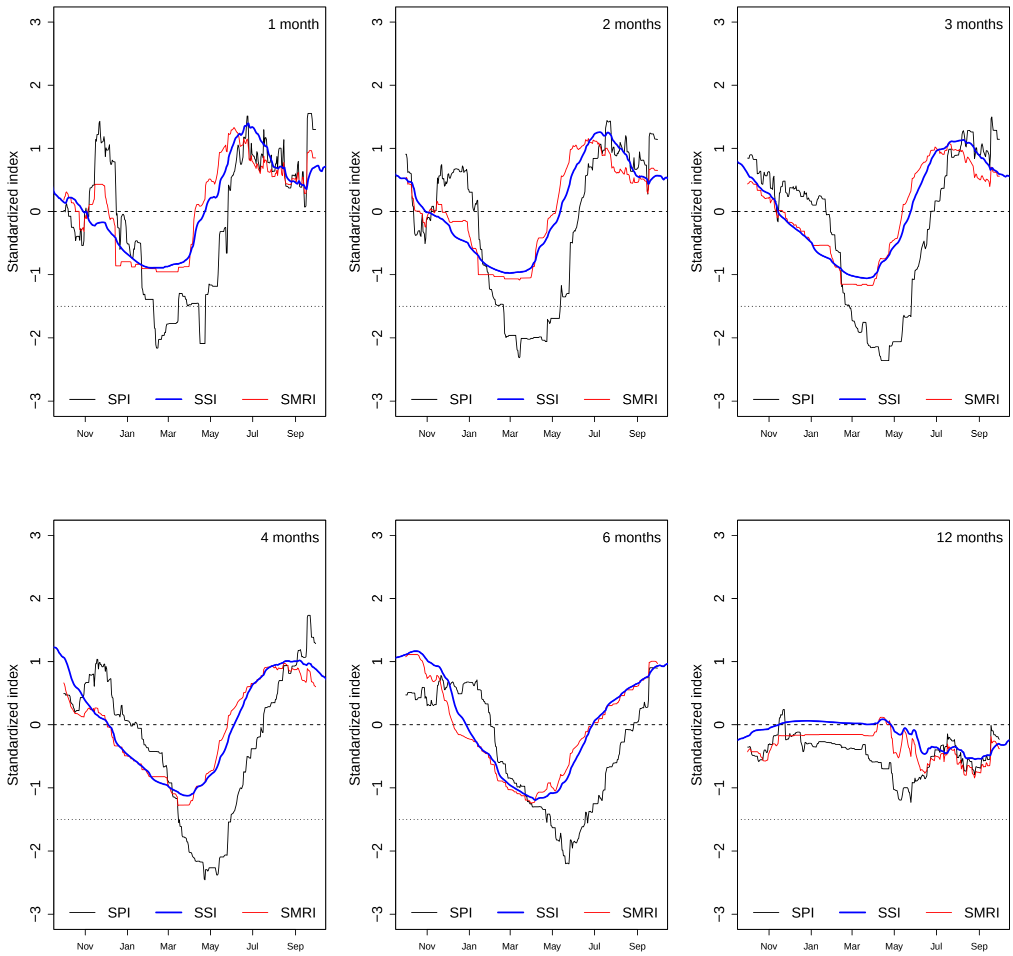

Figure 3.5: Standardized precipitation (SPI) (black), streamflow (SSI) (blue) and ensemble mean of the snow melt rain index (SMRI) (red) in daily resolution for six different accumulation periods during the spring drought 2011 for the nival Ova da Cluozza catchment.

dex. In the mountainous regions of interest in this study the process first modeled is the delayed storage and release of snow. As for similar approaches such as the SRI, which used a hydrological model in the first step (Shukla and Wood, 2008) or the Standardized Precipitation and Evapotranspiration Index (SPEI), which uses an evapotranspiration estimation in the first step (Vicente-Serrano et al., 2010; Vicente-Serrano et al., 2012) this two-step process means that the resulting index has multiple sources of uncertainty. The most important sources of uncertainty from the snow model are the parameterization of the degree-day model, the spatial discretization of elevation as well as data uncertainty. Model parameterization and spatial discretization were addressed by ensemble approaches using parameterizations stemming from no prior knowledge, regional knowledge and specific catchment knowledge. The calculation of the actual index is then influenced by semi-objective decisions including that for a theoretical distribution function and finally the choice for an aggregation period to be used.

Table 3.2: Mean values of F_{bench} for different aggregation periods and all catchments.

Aggregation time [months]	Mentue	Sense	Sitter	Riale di Calneggia	Allenbach	Ova da Cluozza	Dischma
Full period							
1	0.008	0.079	0.197	0.341	0.332	0.573	0.465
2	0.010	0.093	0.226	0.366	0.405	0.610	0.479
3	0.011	0.100	0.268	0.396	0.448	0.629	0.488
4	0.012	0.104	0.284	0.424	0.468	0.639	0.492
6	0.012	0.107	0.261	0.438	0.476	0.620	0.475
12	-0.003	0.029	0.073	0.181	0.182	0.333	0.187
Dry periods							
1	0.004	0.059	0.216	0.386	0.296	0.635	0.514
2	0.004	0.065	0.215	0.390	0.361	0.667	0.541
3	0.005	0.064	0.245	0.412	0.393	0.679	0.558
4	0.008	0.067	0.251	0.436	0.412	0.685	0.553
6	0.011	0.083	0.265	0.438	0.453	0.649	0.498
12	-0.009	0.026	0.097	0.192	0.164	0.352	0.193

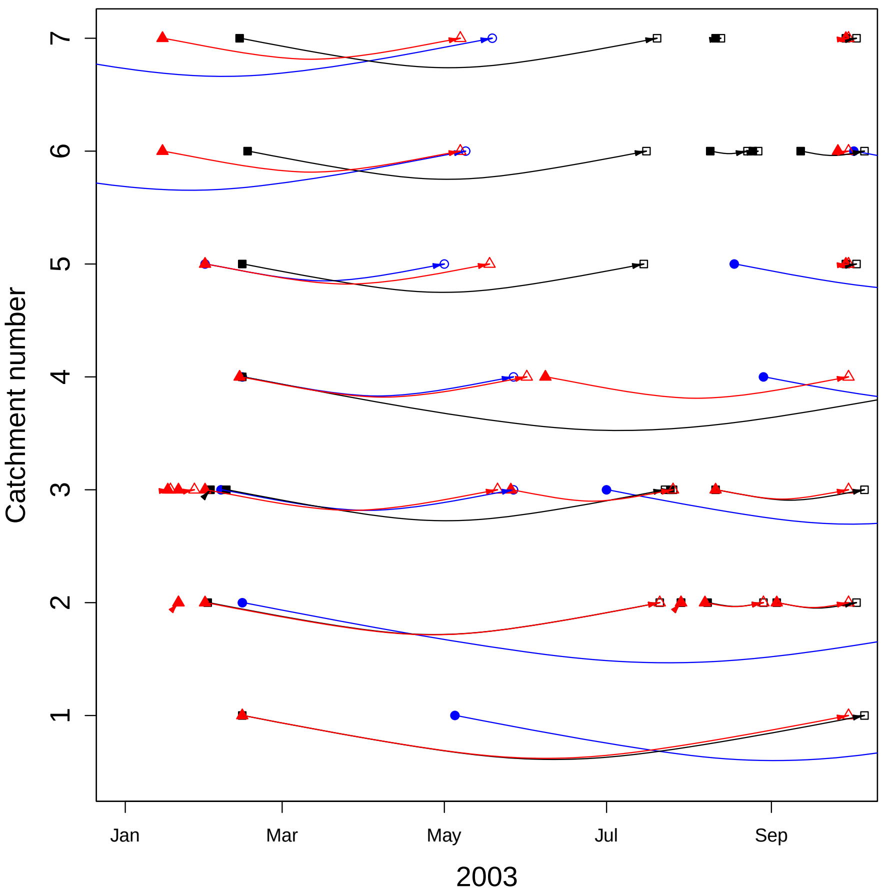

Figure 3.6: Starting dates of the summer drought 2003 (index <0) for standardized precipitation (SPI) (black), streamflow (SSI) (blue) and ensemble mean of the snow melt rain index (SMRI) (red) for the accumulation period of three months.

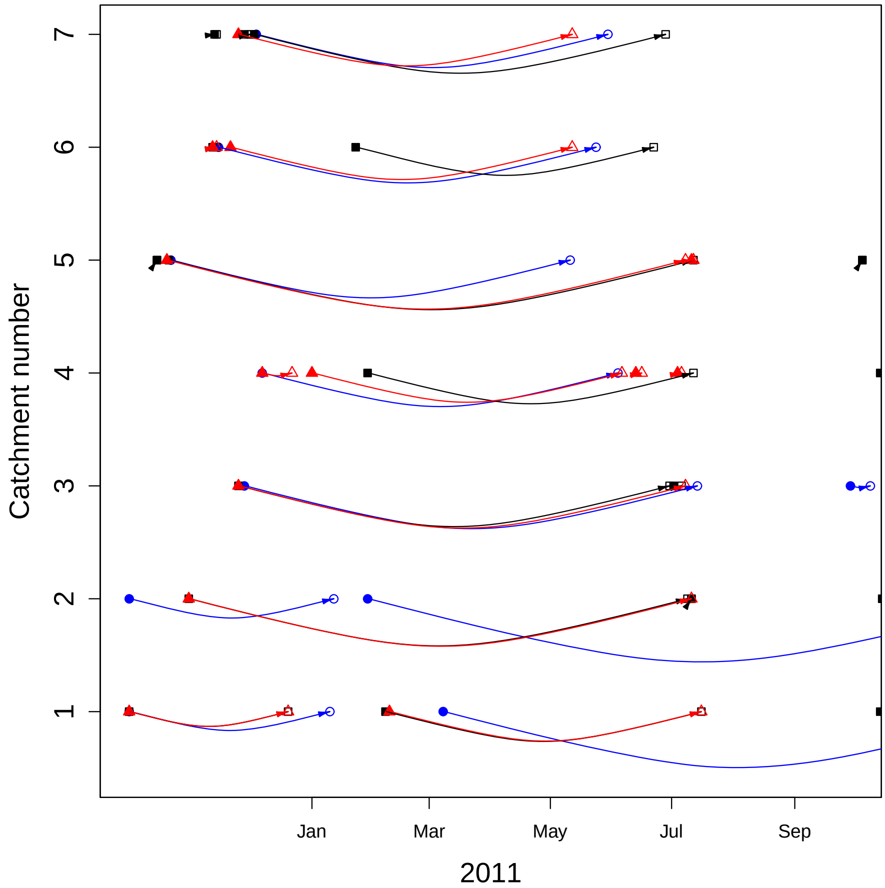

Figure 3.7: Starting dates of the spring drought 2011 (index <0) for standardized precipitation (SPI) (black), streamflow (SSI) (blue) and ensemble mean of the snow melt rain index (SMRI) (red) for the accumulation period of three months.

The Monte Carlo approach that was used is a common way to test the sensitivity to model parameterization (e.g., Demaria et al., 2007). The results showed variation in the performance of the SMRI. However, for the snow influenced catchments and for most parameter combinations, the entire parameter range resulted in an SMRI that was much closer to a hydrological drought description than the SPI for both the entire observation period as well as for the dry periods only. For the catchments with less snow influence there is no disadvantage compared to the SPI. Increasing the knowledge about the snow model parameters of a catchment decreased the uncertainty. However, there was not an increase but a slight decrease of the performance found. This decrease is counter-intuitive but might be explained by the fact that the prior knowledge parameters were derived by calibration of a full hydrological model. The optimal snow parameter values derived in this way might be model specific and not be those providing best results for the SMRI, when soil or groundwater were not considered. These results indicate that the use of an ensemble of random parameters actually might be the most appropriate approach after all. Overall, the parameterization of the snow model has only a minor influence on the systematic performance of the SMRI. Propagating an ensemble is generally a useful way to illustrate the degree of uncertainty that is associated with a model simulation (Pappenberger and Beven, 2004; Montanari, 2005; Choi and Beven, 2007). An ensemble creates more robust results, that depend less on a choice of the parameter values.

Using elevation zones in the melt model instead of a lumped mean elevation, improved the performance of the SMRI. The resulting reduction of the range of SMRI values can be attributed to the explicit consideration of higher elevations. Here, the influence of the threshold temperature (see Appendix) is smaller, thus causing a more stable snow cover and hence less variability in modeled snow melt. Still, when no information on the elevation distribution is available the simpler approach of using the catchment mean elevation resulted in values of F_{bench} greater than zero, meaning the SMRI is closer to the SSI than the SPI. Other studies have found the use of only one lumped elevation zone to result in poor runoff simulations (Uhlenbrook et al., 1999), the aggregation over at least a month in this study compensates the errors as the main effect if different elevation zones is a shift in the timing of snow melt.

Finally, the choice of the climate data input will affect the results. The snow model was driven with uncorrected precipitation data as corrected precipitation data is not a standard data product in Switzerland. Thus there can be errors due to precipitation undercatch - especially in winter when precipitation falls as snow (Rasmussen et al., 2012). However, the resulting bias affects the calculation of both indices, the SMRI and the SPI. Hence, the comparison of the indices and the results presented in this study should not be affected. The grid product used is the result of a well-validated interpolation from a dense network of climate stations. In other mountain regions of the world with less dense networks, interpolation will be more challenging and the errors may be higher.

The choice to include just one key process, i.e., snow accumulation and melt, in the model used to compute the SMRI has implications for the seasonal performance of the new index. The measure of comparison decreases in the months of May to August for the strongly snow influenced catchments. These are the months with the highest evapotranspiration, a process that was not modeled here, but could be considered in a similar way to the SPEI approach (Vicente-Serrano et al., 2010) or in a full hydrological model using the SRI approach (Shukla and Wood, 2008). For many snow-dominated catchments, including those used in this study, the performance gain by including processes other than snow is expected to be small. Despite the exclusion of evapotranspiration, over the entire year the SMRI was closer to the SSI than the SPI and particularly so in the months of snow melt. For the catchments with a pluvial regime, the difference between SPI and SMRI as an indicator for streamflow drought conditions is small or not existent.

There has been some debate over the general concept of standardization which includes fitting a distribution to heavily skewed hydro-meteorological data rather than using empirical percentiles. Empirical percentiles have been used mostly in studies that extract further drought characteristics below a threshold to define severity-area-duration or frequencies (and return periods) (e.g., Cancelliere and Salas, 2010; van Vliet et al., 2012). The concept of standardization has been used mostly for the analysis of entire time series and the propagation of drought through the hydrological cycle (e.g., Hayes et al., 1999; Shukla and Wood, 2008). In this study we chose the SPI approach for consistency and comparability to currently used drought monitoring and early warning efforts. Even though Vicente-Serrano et al. (2011) found differences in mean, standard deviation and in the estimation of extreme quantiles for the different distributions, the major dry and moist episodes, regardless of which distribution function was used, were clearly identified.

4.2 Application potential

Similar to other existing drought indices, the new index can be calculated for different aggregation periods. The co-evolution of SPI, SSI and the new SMRI during two recent drought events showed that with an increasing aggregation period, the SMRI and SPI approximate each other for the studied catchments. The SMRI is thus considered useful to indicate streamflow droughts, that occur in humid to semi-arid mountain regions on a time scale below one year due to the seasonal character of snow storage. The SMRI seems especially suited for warm snow season droughts (Van Loon and Van Lanen, 2012) as the one in spring 2011.

The slightly greater performance difference between SMRI$_{elev}$-1 and SMRI-1 compared to SMRI$_{elev}$-3 and SMRI-3, especially in the catchments with an elevation range between 1000 and 2000 m a.s.l., can be explained with different phases of melt and accumulation that occur in the different elevation zones of a catchment. These differences matter less on longer aggregation time scales as net snow melt amounts for different elevation zones converge.

In the temperate humid climate of Switzerland, snow melt and precipitation occur in the same season, as rainfall is uniformly distributed over the year. This requires shorter aggregation periods to be considered for the calculation of the indices than in a Mediterranean climate, where wet and dry seasons are clearly separated. Where such a clear separation does not exist, other indices that include end-of season snow pack directly as, for example, the SWSI, will be less useful for drought assessment.

Ideally, an index also needs to be suitable for regional comparisons, i.e., easily applicable with distributed or gridded climate datasets and without further information needs. The SWSI, for instance, requires information on the different contributions of precipitation, snow, runoff and reservoir storage as well as their elevational, seasonal and inter annual variations. As Shukla and Wood (2008) stated, a runoff index complements the SPI and can serve to understand the actual hydrological situation concerning droughts. The SRI includes all runoff generation processes including snow melt in the modeled runoff. The strength of a runoff-based index is that it can be used for forecasting and is sensitive to hydrologic initial conditions such as snow conditions in spring (Shukla and Wood, 2008). However, simulated grid runoff cannot be validated. Validation of the SMRI approach with SSI from streamflow observations in meso-scale catchments across a gradient of increasing snow influence, as proposed in this study, shows that for these cases the simpler approach is a suitable alternative to describe the evolution of hydrological drought situations.

5 Conclusions

The SMRI, as introduced in this study, combines the low data requirements of the SPI with the explicit consideration of snow accumulation and melt. The analyses of the new index demonstrates its usefulness to indicate hydrological droughts in snow influenced catchments, with specific advantages in those climatic regions where snow melt and rainy

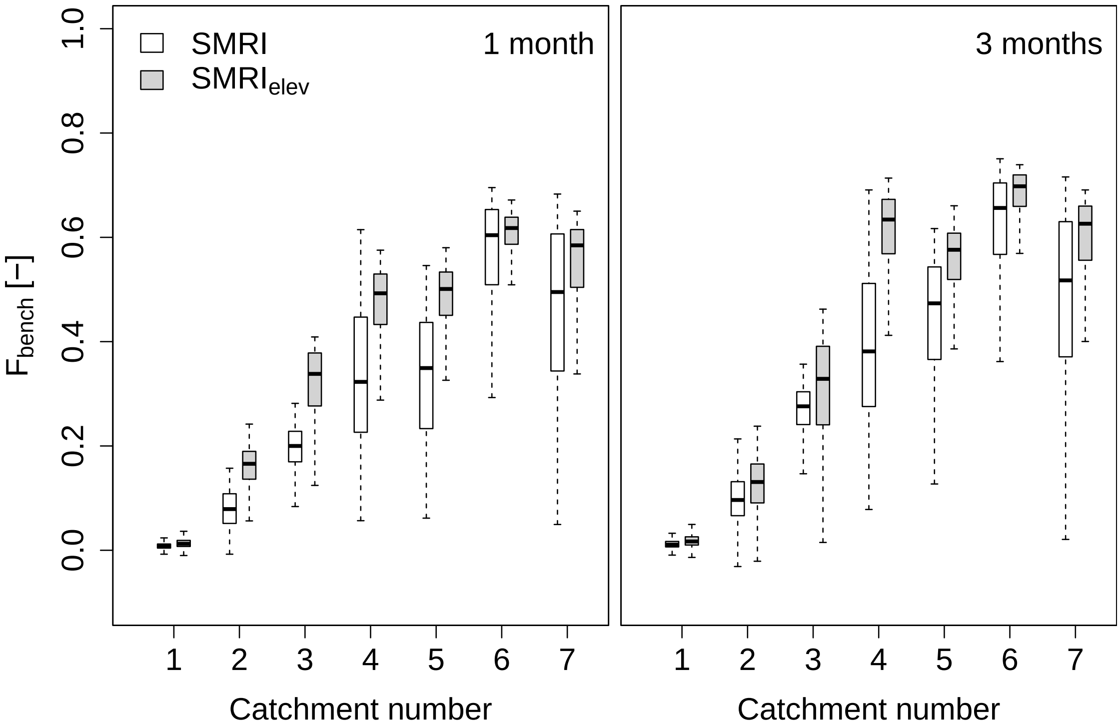

Figure 3.8: Comparison of the distribution of the measure of comparison for *SMRI* and $SMRI_{elev}$ for the aggregation periods of one (left) and three (right) months. Boxplots as in Figure 3.1.

season coincide. This case study with Swiss catchments suggests a closer description of hydrological droughts by the SMRI than by SPI. Following the gradient of snow influence, the more a catchment is influenced by snow the more worthwhile it is to complement the SPI with the SMRI.

The SMRI is a somewhat more complex index than the SPI as it also uses temperature data to consider snow processes in the computation. Thus it has some additional sources of uncertainty. The aggregation period can be adjusted to the typical seasonality of the hydrological regime, water resources use and management requirements. As the index corresponds to the SPI during seasons or years without snow, it can be used without problems for drought monitoring and assessment over diverse mountain regions with regime transitions.

Despite the different realizations derived from different parameter sets of the snow model, the SMRI described both the hydrological situation in general as well as dry periods in particular closer than the SPI particularly in snow influenced catchments. We therefore recommend using the SMRI for drought monitoring in snow influenced catchments without streamflow measurements.

Appendix - Snow model

The new SMRI was based on snow melt computations using a simple degree-day snow model. Whenever the observed air temperature (T) [°C] is lower than a threshold temperature (T_T) [°C] precipitation is added to the snow storage (accumulation A [mm]). In addition to the accumulation the liquid water content S_{liquid} [mm] in the snow pack is also calculated. S_{liquid} is calculated accounting for precipitation (P) [mm], melt M [mm] and refreezing (R) [mm] and has an upper bound constrained by the water holding capacity (C_{WH})[-]. Refreezing (R) is determined by S_{liquid} of the day before, a degree-day factor C_M [mm/day°C] and a refreezing factor C_{FR} [-]. Melt is constrained by the preceding accumulation and calculated using C_M, T_T and T. The contribution to surface runoff Q [mm] is all water that exceeds C_{WH} of the snow pack. For this study C_{WH} was kept constant at a value of 0.1. (see pseudo code below)

---

if $T(t) < T_T$ **then**
  $R(t) = min(S_{liquid}(t-1), C_{FR} * C_M * (T_T - T(t)))$
  $A(t) = A(t-1) + P(t) + R(t)$
  $S_{liquid}(t) = S_{liquid}(t-1) - R(t)$
else
  $M(t) = min(A(t-1), C_M * (T(t) - T_T))$
  $A(t) = A(t-1) - M(t)$
  $S_{liquid}(t) = S_{liquid}(t-1) + P(t) + M(t)$
  if $S_{liquid}(t) > C_{WH} * A(t)$ **then**
    $Q(t) = S_{liquid}(t) - C_{WH} * A(t)$
    $S_{liquid}(t) = C_{WH} * A(t)$
  end if
end if

APPENDIX 4

Comparison of hydrological model structures based on recession and low flow simulations*

Abstract

Low flows are often poorly reproduced by commonly used hydrological models, which are traditionally designed to meet peak flow situations. Hence, there is a need to improve hydrological models for low flow prediction. This study assessed the impact of model structure on low flow simulations and recession behaviour using the Framework for Understanding Structural Errors (FUSE). FUSE identifies the set of subjective decisions made when building a hydrological model and provides multiple options for each modeling decision. Altogether 79 models were created and applied to simulate stream flows in the snow dominated headwater catchment Narsjø in Norway (119 km^2). All models were calibrated using an automatic optimisation method. The results showed that simulations of summer low flows were poorer than simulations of winter low flows, reflecting the importance of different hydrological processes. The model structure influencing winter low flow simulations is the lower layer architecture, whereas various model structures were identified to influence model performance during summer.

1 Motivation

Hydrological low flow periods and droughts affect water supply for drinking water, irrigation, industrial needs, hydropower production and ecosystems. Their occurrence is also of importance regarding environmental flow and water quality requirements, which are strongly connected to critical low flows (Vogel and Fennessey, 1995). Low flow and droughts affect many sectors and occur in every country albeit in different perceived

*Published as: Staudinger, M., Stahl, K., Seibert, J., Clark, M. P., and Tallaksen, L. M. (2011). Comparison of hydrological model structures based on recession and low flow simulations. Hydrology and Earth System Sciences, 15(11), 3447-3459.

severity. There is a wide range of consequences related to low flow and drought and monitoring and modelling of low flow are crucial for their analysis and prediction. However, low flows are poorly reproduced by many hydrological models since these are traditionally designed to simulate the runoff response to rainfall.

A revision of model concepts regarding low flows requires a clear understanding of the model's structural deficits; in other words "when does it go wrong and which part of the model is the origin?" (Reusser et al., 2009). A common approach to investigate the impact of the differences in model structure is to perform model intercomparison experiments, e.g. Henderson-Sellers et al. (1993), Reed et al. (2004), Duan et al. (2006), Breuer et al. (2009) and Holländer et al. (2009). Such experiments have been helpful to explore model simulation performance of lumped (Duan et al., 2006; Breuer et al., 2009), semi-distributed (Duan et al., 2006; Holländer et al., 2009) and distributed (Henderson-Sellers et al., 1993; Reed et al., 2004; Holländer et al., 2009) models in a consistent way using the same input data. The reasons for the differences, however, remain unclear since each model uses different interacting parametrisations to simulate the hydrological processes (Clark et al., 2008). Perrin et al. (2001) studied the relation between the number of optimized parameters and model performance in a multi-model, multi-catchment experiment, and discussed the problem of over-parametrisation and parameter uncertainty.

Discrepancies between observed and simulated streamflow can arise from errors in the input data rather than weaknesses in model structure. This complicates the investigation of the impact of the differences in model structure. Clark et al. (2008) created a computational framework that enables a separate evaluation of each model component. The Framework for Understanding Structural Errors (FUSE) differs from others as it modularises individual flux equations instead of linking available submodels. FUSE identifies the set of subjective decisions while creating a hydrological model and offers multiple options for each model decision. This approach can thus help to get a better understanding of the hydrological processes occurring. Clark et al. (2008) first introduced FUSE, as a diagnostic tool to evaluate the performance of hydrological model structures using the Nash-Sutcliffe efficiency for two climatically different catchments. Clark and Kavetski (2010) evaluated several classes of numerical time stepping schemes in order to find appropriate numerical methods used to solve the governing model equations of hydrological models. The experimental setup included beside different distinct time stepping algorithms, eight conceptual rainfall runoff models derived from the parent models. Another recent application of FUSE is documented in the two-part series of McMillan et al. (2011) and Clark et al. (2011). First, they used precipitation, soil moisture and streamflow data to estimate the dominant hydrological processes of a catchment. Then, plausible representations of these processes in conceptual models were formulated (McMillan et al., 2011). In the second part, they evaluated FUSE models regarding their capability to simulate those processes Clark et al. (2011).

Commonly, streamflow recession is modelled as the outflow from a, or a set of, linear or non-linear reservoirs. In periods with no input, i.e. precipitation or snow melt, outflow from the reservoirs control the streamflow and thus, the model behaviour during low flow. Real hydrological processes can be more complex. Therefore, it is of interest to have a closer look at the hydrograph recession, and carefully evaluate model simulations of recession behaviour. The shape of the observed recession curve reflects the gradual depletion of water stored in a catchment during periods with little or no precipitation.

Initially, the recession curve is steep as quick flow components like overland flow and subsurface flow contribute to streamflow. The recession curve flattens with time as e.g. delayed water from deeper subsurface storages contributes, and may become nearly constant if sustained by outflow from the groundwater storage or from a glacier (Smakhtin, 2001). The recession curve describes in an integrated manner how different factors in a catchment influence the generation of streamflow in dry weather periods (Tallaksen, 1995). Hydrogeology, relief and climate have been found to be the most important catchment properties affecting the recession rate (Tallaksen, 1995). Catchments with a slow recession rate are typically groundwater dominated, while impermeable catchments with little storage show faster recession rates. Moreover, summer recessions are usually faster than autumn or winter recessions (e.g. Federer, 1973; Tallaksen, 1995).
Several studies exist that link recession analysis with the structure of hydrological models (e.g. Ambroise et al., 1996; Wittenberg, 1999; Clark et al., 2009; Harman et al., 2009). In this study the model structures are systematically analysed using FUSE. The associated model performance is evaluated with respect to the ability to simulate low flows and recession behaviour. This is done for one catchment only to allow a more detailed insight in the model structures. The main objective is to investigate the relative influence of a single model structure on the model performance. As there are distinct differences in the recession rates found for summer and winter, one task is to study how model structure is connected to the seasonal performance for low flow simulation. This paper aims to contribute to the improvement of hydrological models for low flow prediction.

2 Data and study area

The data are from the 119 km^2 headwater catchment Narsjø, located in the South-East of Norway (Fig. 4.1) with an altitude range between 737 and 1595 m a.s.l. (Engeland, 2002). Narsjø is a subcatchment of the Upper Glomma basin, which is characterised by a continental climate with cold winters and relatively warm summers (Engeland, 2002). The annual snow melt flood dominates the hydrological regime. The most pronounced low flow period occurs in winter, caused by precipitation being stored in snow and ice. A second low flow period occurs in summer, caused by a lack of precipitation and losses due to evapotranspiration (Engeland, 2002). The geology can be divided into two main areas: one area consists of schists and phyllites that occur in combination with fine grained till soil, the other area consists of igneous rocks (granite, gneiss and gabbro) usually in combination with coarser till (Engeland, 2002). This geological characteristic influences the properties of soil and vegetation. The quaternary remains, consisting of several types of till and fluvial deposits as well as bogs and lakes, form a wide, open mountain landscape with gentle slopes. The land cover is barely influenced by humans (0.3 % agricultural land) and is composed of 23.7 % forest, 60.9 % open land, 12.0 % bogs and 3.0 % lakes (Engeland, 2002).
The streamflow data used are daily time series of observed discharge measured at the outlet of the Narsjø catchment (provided by the Norwegian Water Resources and Energy directorate, NVE). In addition, daily time series of precipitation interpolated from 12 surrounding meteorological stations and potential evaporation (Beldring et al., 2003) were available. The time series cover the period from 6 May 1981 to 31 December 1995.

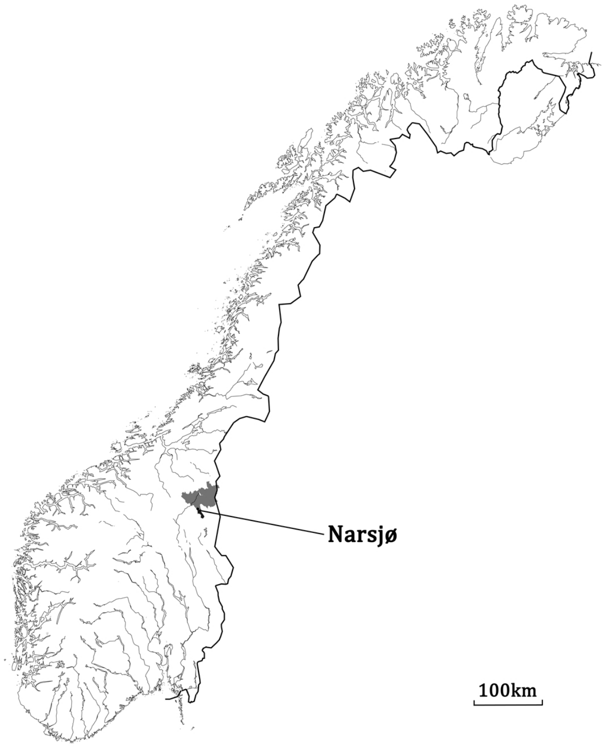

Figure 4.1: Location of the Narsjø catchment (modified after Beldring et al., 2003).

3 Methods

3.1 Snow accumulation and melt

Narsjø is a snow dominated catchment, however, there was no snow routine implemented in the version of FUSE used for this study. Hence, the input data was pre-processed with a snow accumulation and melt model. This corresponds to an implemented snow routine. Here, a simple degree day method was applied. The daily change in snow water equivalent ΔSWE [mm day^{-1}] is equal to the difference in the daily snow accumulation a_s [mm day^{-1}] and the daily snow melt m_s [mm day^{-1}] (Eq. 4.1).

$$\Delta SWE = a_s - m_s. \tag{4.1}$$

The snow model separates the precipitation P [mm day^{-1}] into rain and snow using a temperature threshold. Hence, there is only snow accumulation a_s in the catchment when the measured temperature T [$^\circ C$] is below the threshold temperature T_{acc} (Eq. 4.2).

$$a_s = \begin{cases} 0, & T \geq T_{acc}, \\ P, & T < T_{acc}. \end{cases} \tag{4.2}$$

In this study T_{acc} was set to 1.0 $°C$. The daily snow melt m_s was computed (Eq. 4.3) with a melt factor M_f of 3.0 $°C^{-1}$ day^{-1} and a melt threshold temperature T_{melt} of $0°C$.

$$m_s = \begin{cases} M_f(T - T_{melt}), & T \geq T_{melt} \quad \text{and} \quad SWE > 0, \\ 0, & T < T_{melt} \quad \text{and} \quad SWE = 0. \end{cases} \tag{4.3}$$

The chosen melt factor was based on Seibert (1999) who found melt factors in Sweden to vary between 1.5 and 4 mm $°$ C^{-1} day^{-1}, where the first value is suited for open and the latter for forested sites. The degree day method was extended with a refreeze factor r_f [-] which accounts for rain that does not directly contribute to runoff due to the water holding capacity of an existing snow cover (Eq. 4.4).

$$P = \begin{cases} 0, & T \geq T_{acc}, \\ P, & T \geq T_{acc} \quad \text{and} \quad P \geq r_f \quad SWE, \\ (1 - r_f)m_s, & T \geq T_{acc} \quad \text{and} \quad P < r_f \quad SWE. \end{cases} \tag{4.4}$$

3.2 FUSE framework

The use of FUSE as a diagnostic tool to detect the impact of model structure involved the following three steps: (1) prescription of the type of model (2) definition of the major model-building decisions and (3) preparation of multiple options for each model building decision (Clark et al., 2008). In this study, the type of model was limited to lumped hydrological, that were run at a daily time step (although the models are not limited to a daily time step). Four conceptual parent models were selected to be recombined to new FUSE-models: ARNO-VIC (Zhao, 1977), TOPMODEL (Beven and Kirkby, 1979), PRMS (Leavesley et al., 1983) and SACRAMENTO (Burnash, 1995). Simplified wiring diagrams of the generating parent models are shown in Figure 4.2. The selection of the parent FUSE models was here limited to four well known models, covering common principles used in conceptual hydrological models. All parent models consist of equally plausible structures and the important processes could be broken down into fluxes occurring in the upper layer and lower layer, evaporation, percolation, subsurface flow and surface runoff (model building options).
Some processes were not explicitly modelled, including interception by the vegetation canopy as well as specific surface energy balance calculations. Routing was calculated by a two parameter Gamma distribution (Press et al., 1992). Thus, all models represent the subsurface with a similar level of detail and thus differences that emerged from different plausible model structures were emphasised rather than differences due to the set of processes represented. The model decision options that were made separately for each of the FUSE models are described next (more details to the decision options e.g. equations can be found in Clark et al. (2011)). A summary of those decisions that were permuted for this study can be found in Table 4.1 and the abbreviations from Table 4.1 will be referred to later in the text.

Upper layer The water content of the upper soil layer was either defined as a single state variable or split into tension storage and free storage, with an additional option to further subdivide the free storage into below and above field capacity (Table 4.1).

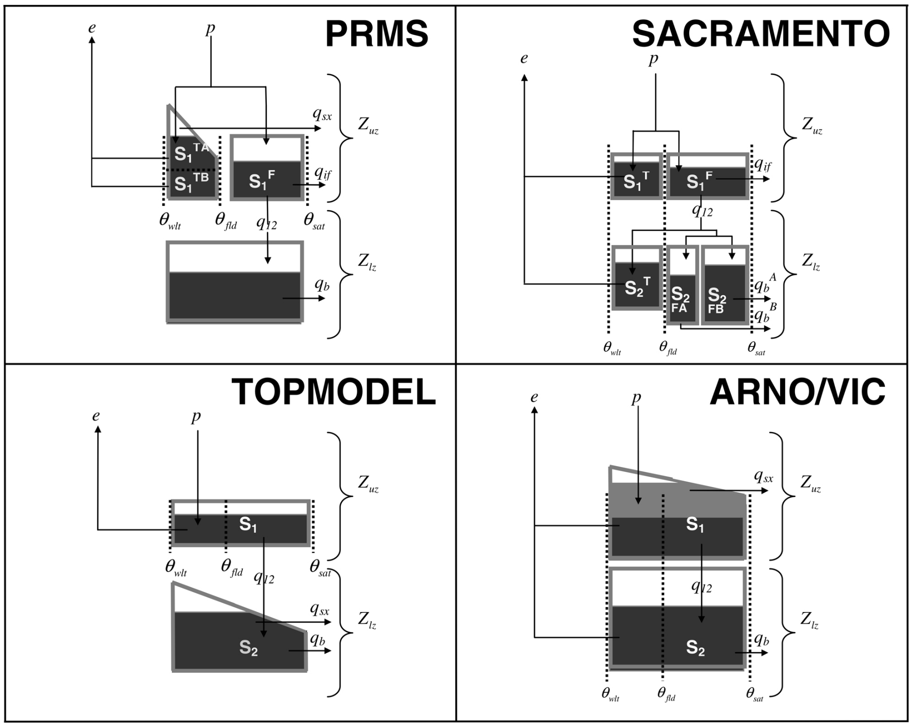

Figure 4.2: Simplified wiring diagrams of the parent models (from Clark et al., 2008).

Lower layer The lower soil layer was either defined by a single state variable with unlimited storage and no lower layer evapotranspiration, by a single state variable with fixed storage and no lower layer evapotranspiration or as a tension storage combined with two parallel tanks (Table 4.1). All subsurface flow options (see below) are closely connected to the lower layer, this is why the choice of subsurface flow and lower-layer option is realised as a single model decision within FUSE (Clark et al., 2008).

Evaporation Evaporation was parameterised by the sequential evaporation scheme (Clark et al., 2008): first potential evaporative demand is supplied by evaporation from the upper layer and then any residual demand by water from the lower layer.

Percolation In FUSE there are three percolation options each having two parameters (Table 4.1). The architecture of the parent model VIC is equivalent to the gravity drainage term in the Richard's equation (e.g. Boone and Wetzel, 1996), often resulting in a large exponent to limit drainage below field capacity (water can percolate from the wilting point to saturation). The equation used in PRMS does not allow drainage below field capacity (water can percolate from the field capacity to saturation). Non-linearities in the SACRAMENTO parametrisation are controlled by the moisture content in the lower layer, meaning percolation will be fastest when the lower layer is dry (Clark et al., 2008). All three options were used as model decision options.

Table 4.1: FUSE model decision options.

Model structure	Model option	Abbreviation
Upper layer architecture U	Upper layer divided into tension and free storage	$U_{tension1}$
	Free storage plus tension storage sub-divided into recharge and excess	$U_{tension2}$
	Upper layer defined by a single state variable	$U_{onestate}$
Lower layer architecture and subsurface flow L	Tension storage combined with two parallel tanks	$L_{tens2pll}$
	Storage of unlimited size combined with linear fraction rate	$L_{unlimfrc}$
	Storage of unlimited size combined with power recession	$L_{unlimpow}$
	Storage of fixed size with non-linear storage function	$L_{fixedsiz}$
Surface runoff S	ARNO/Xzang/VIC parametrisation	$S_{arno/vic}$
	PRMS variant; fraction of upper tension storage	S_{prms}
	TOPMODEL parametrisation	S_{tmdl}
Percolation P	Water from field capacity to saturation available for percolation	P_{f2sat}
	Water from wilting point to saturation available for percolation	P_{w2sat}
	Percolation defined by moisture content in lower layer architecture	P_{lower}

Subsurface flow There are four subsurface flow options (Table 4.1). Subsurface flow was modelled either by a single linear storage, by two parallel connected linear reservoirs or by nonlinear storage functions like in ARNO/VIC or TOPMODEL (Clark et al., 2008). TOPMODEL requires a distribution of topographic index values for each catchment (Beven and Kirkby, 1979). For the Narsjø catchment the distribution was derived using a three-parameter Gamma distribution following Sivapalan et al. (1987).

Surface runoff Surface runoff was generated using a saturation-excess mechanism, when it rains on saturated areas of the basin. The surface runoff is distributed according to the topographic index distribution (defined in Clark et al., 2008).

Bucket overflow Additional fluxes of water may occur when one of the storages reaches its capacity. In the upper layer, the bucket overflow from the primary tension storage carries over precipitation that falls into the second tension storage. The bucket overflow from a tension storage carries precipitation into a free storage and from the free storage it adds to surface runoff. In the lower soil layer, the bucket overflow from tension storage forms additional percolation into free storage and from free storage again additional subsurface flow. Following Kavetski and Kuczera (2007), logistic functions were used to smooth the thresholds associated with a fixed capacity of model storages.

Routing The time delay in runoff was modelled using a two-parameter Gamma distribution (Press et al., 1992), with an adjustable mean of the Gamma distribution. The shape of the time delay histogram, however, was fixed by setting the shape parameter to 3.0 to keep the number of adjustable parameters small.

3.3 Model calibration

All FUSE models were calibrated using the Shuffled Complex Evolution algorithm (SCE) which was parameterised based on the recommendations of Duan et al. (1994). A maximum of 10 000 trials was allowed before the optimisation was terminated. Within five shuffling loops the value had to change by 10 % or the optimisation was terminated. The number of complexes in the initial population was set to 10. Each complex contained $2N_{opt} + 1$ points, each sub-complex $N_{opt} + 1$ points and $2N_{opt} + 1$ evolution steps were allowed for each complex before shuffling, where N_{opt} was the number of parameters to be optimised in the calibration procedure, respectively. The algorithm was used to minimise the mean absolute relative error (F_{MARE}) (Eq. 4.5). F_{MARE} was chosen as objective function, because it emphasises low to medium flows. MARE ranges between zero and infinity with the optimum at zero. Calibration and validation by the two objective functions is done on the entire series, but both objective functions chosen emphasize the lower flow ranges of the hydrograph.

$$F_{MARE} = \frac{1}{n} \sum_{i=1}^{n} \frac{|Q_{obs}(i) - Q_{sim}(i)|}{Q_{obs}(i)} \tag{4.5}$$

The calibration was performed for 15 yr using a three years spin up period. As recommended by Clark and Kavetski (2010) for conceptual hydrological models, the fixed step implicit Euler method was used as numerical time stepping scheme.

3.4 Low flow and recession analysis

The performance of the model was then evaluated using the logarithmic Nash-Sutcliffe efficiency. F_{logNSE} that also emphasizes model performance in the low flow range. F_{logNSE} was based on log-transformed streamflow series from observation Q_{obs} and simulation Q_{sim} (Eq. 4.6). This metric ranges between minus infinity and one and a perfect model would result in 1.

$$F_{logNSE} = 1 - \frac{\sum_{i=1}^{n} (\ln(Q_{obs}(i)) - \ln(Q_{sim}(i)))^2}{\sum_{i=1}^{n} (\ln(Q_{obs}(i)) - \ln(\bar{Q}_{obs}))^2} \tag{4.6}$$

As a good model should be able to produce reasonable results for a range of objective functions and not only for the one it was calibrated to, the performance was evaluated using F_{logNSE}, whereas the models were calibrated using F_{MARE}.
Several studies use recession analysis to infer the exponent in a non-linear storage (Ambroise et al., 1996; Wittenberg, 1999; Clark et al., 2009; Kirchner, 2009), or, more generally, provide guidance on the structure of a hydrological model (Clark et al., 2009;

Harman et al., 2009). Recession analysis is also useful as a diagnostic tool for model evaluation (McMillan et al., 2011; Clark et al., 2011). In this study the relationship between the negative change in streamflow over time $-\frac{dQ}{dt}$ [mm day^{-2}] and the corresponding streamflow Q [mm] was analysed (Brutsaert and Nieber, 1977). For the evaluation of the model performance of recessions both modelled and observed data were used. The method was modified by using flexible (instead of fixed) time steps scaled to the observed streamflow ΔQ between time steps as recommended by Rupp and Selker (2006). Our study is based on daily observations and similar to Palmroth et al. (2010), the lower and upper limits of the time step are set to 1 and 5 days, respectively. The time step is then found by setting the maximum difference in ΔQ (threshold) between to time steps equal to 0.1 % of the mean observed streamflow at that point. As both $-\frac{dQ}{dt}$ and Q span several orders of magnitude, their relation is plotted in log-log-space. The data points in the plots including all recessions of the hydrograph and might thus be composed of both subsurface and overland flow. Overland flow would mainly affect the upper range of streamflow values. Hence, the upper range in the plots of $-\frac{dQ}{dt}$ and Q should be treated with special care if interpreted regarding storage release. In case of an exponential recession (simple linear storage model) the relation can be expressed as in Eq. (4.7), where p is a constant. However, a power function results in Eq. (4.8), with the additional coefficient q.

$$\frac{dQ}{dt} = -p\ Q \tag{4.7}$$

$$\frac{dQ}{dt} = -p\ Q^q \tag{4.8}$$

The $-\frac{dQ}{dt}$ versus Q plots can become noisy. Therefore, points in a certain range of Q were averaged to one value representative for this range (binned). Then, a polynomial function was fitted to the relationship between $-\frac{dQ}{dt}$ and Q (Eq. 4.9) (Kirchner, 2009).

$$\ln\left(\frac{-dQ/dt}{Q}\right) \approx a + b\ \ln(Q) + c(\ln(Q))^2 \tag{4.9}$$

The polynomial coefficients were fitted using a least squares regression model. The significance of the regression model was tested with the Kolmogorov-Smirnov goodness-of-fit test (Massey Jr, 1951). The polynomial fitted to the observed recessions is used as a benchmark model (see Seibert, 2001) similar to the mean streamflow being used as a benchmark model for the Nash-Sutcliffe efficiency (NSE). Hence, passing the Kolmogorov-Smirnov test, similar to a NSE above zero, is used as an objective decision for acceptable models (similar or better than the benchmark). The choice of a polynomial follows Kirchner (2009). It is used because of it offers both enough flexibility to adapt to the data and enough smoothness to allow moderate extrapolation beyond the binned relationships. Scatter plots of the coefficients b and c in Eq. (4.9) were then used to compare observed and simulated recession behaviour for the FUSE models that passed the Kolmogorov-Smirnov goodness-of-fit test. The relationship between $-\frac{dQ}{dt}$ and Q is in the following referred to as the "recession relationship".
The recession behaviour was analysed for both the whole year and the individual seasons. The seasonal recessions were derived by splitting the recessions for the whole year into summer and winter recessions. Winter was defined as the time from October 15,

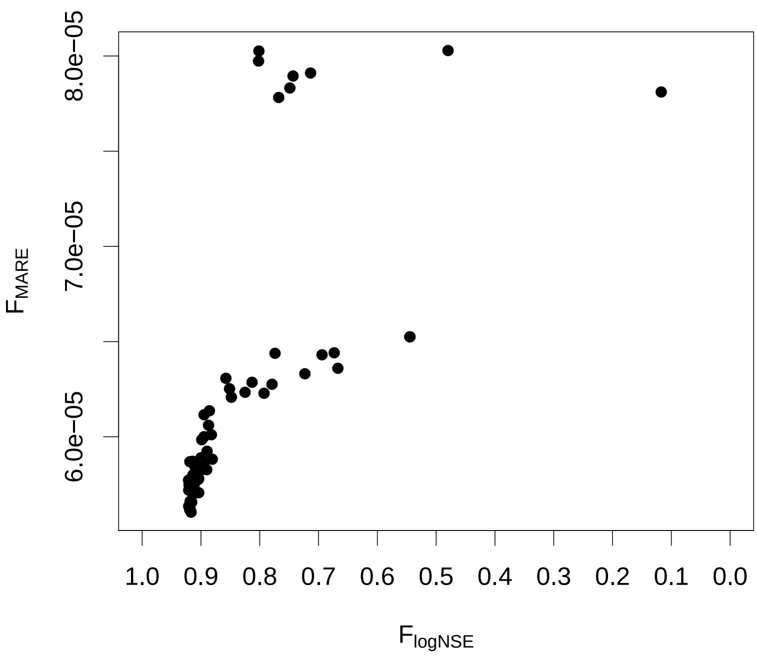

Figure 4.3: F_{logNSE} versus F_{MARE} for the 79 FUSE models after calibration with SCE (Shuffled Complex Evolution algorithm).

when precipitation generally begins to fall as snow in the catchment, to June 15, which is usually towards the end of the snowmelt period.

4 Results

4.1 Calibration

For 73 out of 79 FUSE models the F_{logNSE} was greater than zero. In Fig. 4.3 a scatter plot of the resulting values of the objective functions for both calibration (F_{MARE}) and evaluation (F_{logNSE}) is shown. The axes are ordered from high to low model performance for both measures, which means that the points of best performance group in the lower left corner. It appears that the F_{logNSE} and the F_{MARE} show a similarly good model performance for the F_{logNSE} range from 1 to 0.8. However, for lower F_{logNSE} the two objective functions differ. While the models are considered poorer for F_{logNSE}, F_{MARE} remains at the same level.

4.2 Model performance during low flows

All models with $F_{logNSE} < 0$ used the same combination of lower layer/subsurface flow and percolation options $L_{unlimpow}$ and P_{lower} (see Fig. 4.4). The best models (F_{logNSE} > 0.8) used varying combinations. The majority of the best models, however, used a

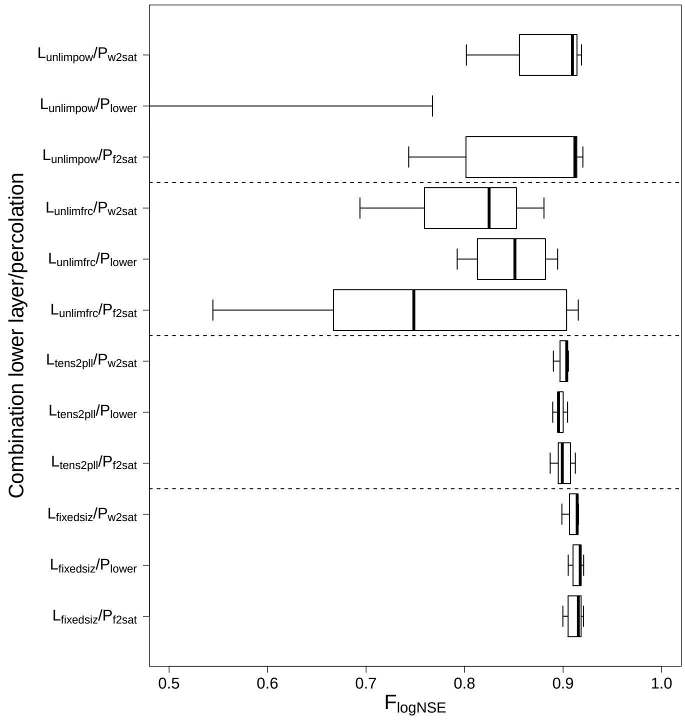

Figure 4.4: Boxplots of the performance of models using different lower layer and percolation combinations. The box of models using $L_{unlimpow}$ and P_{lower} includes model performances (F_{logNSE}) below zero.

lower layer/subsurface flow combination of either $L_{tens2pll}$ or $L_{fixedsiz}$. Many of the poor models used a combination of $L_{unlimfrc}$ for lower layer/subsurface flow and P_{f2sat} for percolation. The poorest models in the group with $F_{logNSE} > 0$ primarily used the same combination of lower layer/subsurface flow and percolation options as found for the poorest performing models ($F_{logNSE} < 0$). All possible upper layer and surface runoff options were found for the poorest performing models.

4.3 Recession behaviour

The observed flow values in the recession periods ranged between 0.2 and 40 mm day^{-1} for Q and between 0.001 and about 15 mm day^{-2} for $-\frac{dQ}{dt}$ and in general showed a linear recession relationship with higher $-\frac{dQ}{dt}$ for higher Q. Most of the modelled recession relationships were similar in range, their shapes, however, differed: some appeared more convex, others more concave and a third group showed nearly a linear recession relationship. In comparison to the observed range, some of the models produced an unrealistic scatter. For example, low flow values were modelled that were below the observed range (Fig. 4.5 f)) and their associated recession slopes were too steep (Fig. 4.5 e) and f)). The latter behaviour was only found for models containing a combination of the lower layer/subsurface flow $L_{unlimpow}$ and the percolation P_{lower}. The model decision options for the example models in Fig. 4.5 are listed in Table 4.2. The combinations including

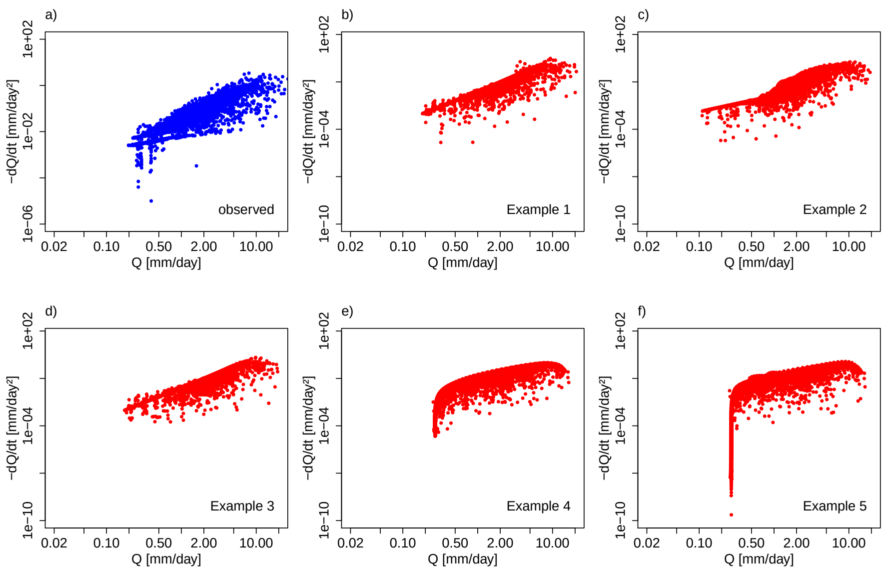

Figure 4.5: Plots of recession relationships for observed recessions (a), blue and five examples of simulated recessions (b–f), red. The model decision options for the examples can be found in Table 4.2.

Table 4.2: Model decision options for the examples in Fig. 4.5.

Example	Upper Layer	Lower Layer	Percolation	Surface runoff
1	$U_{tension2}$	$L_{unlimpow}$	P_{f2sat}	S_{prms}
2	$U_{tension2}$	$L_{unlimfrc}$	P_{f2sat}	S_{tmdl}
3	$U_{tension2}$	$L_{fixedsiz}$	P_{lower}	S_{prms}
4	$U_{tension2}$	$L_{unlimpow}$	P_{lower}	S_{tmdl}
5	$U_{tension2}$	$L_{unlimpow}$	P_{lower}	S_{prms}

the S_{prms} surface runoff option (Fig. 4.5b, d and f) show linear relationships, while the combinations including S_{tmdl} (Fig. 4.5c and e) show convex or concave relationships. Figure 4.5e includes the lower layer/subsurface flow and percolation options $L_{unlimpow}$ and P_{lower} and shows a large range in $-\frac{dQ}{dt}$ for the same flow values. The coefficients b and c from Eq. (4.9) are shown in Fig. 4.6. The b coefficient describes the slope and the c coefficient the curvature of the binned recession relationships. The observation pair can be found at the edge of the group resulting from the simulations having a large b coefficient and a small c coefficient. Most pairs are located in the lower right quarter, i.e. in the area of positive slope and negative curvature. A smaller group can be found for positive b and c coefficients and only few models resulted in negative b and c coefficients. None was fitted with negative slope and positive curvature. The few models that resulted in negative slope and negative curvature used $L_{unlimpow}$ for lower layer and subsurface flow, S_{prms} for surface runoff and P_{lower} for percolation. The models that resulted in both coefficients being positive predominantly used $U_{onestate}$ for the upper layer architecture, often combined with $L_{unlimfrc}$ for lower layer/subsurface

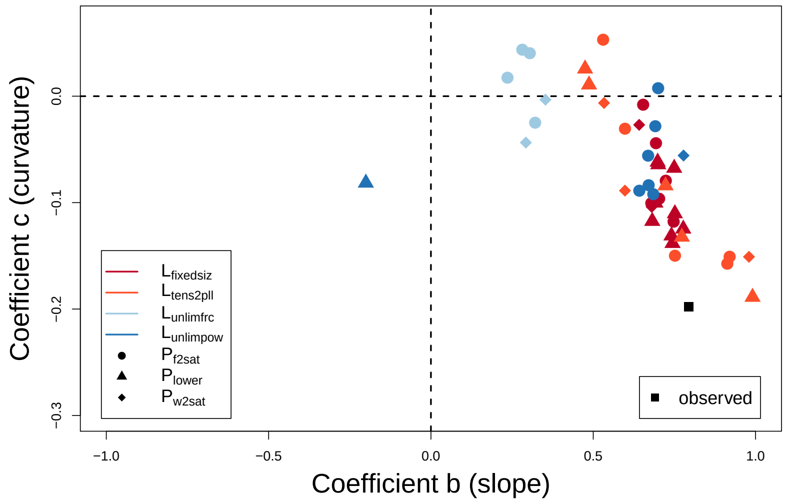

Figure 4.6: Relation between b and c coefficients of the polynomial function fitted to binned recession relationships.

flow. The only differing model decision option for the upper layer architecture within this group was $U_{tension2}$. All surface runoff structure model options were found in this group. However, the S_{tmdl} parametrisation was found only in the particular combination with $U_{onestate}$ and $L_{unlimfrc}$ for the upper and lower layer architecture, respectively. The steepest slopes (coefficient b) were found for models containing the option $L_{tens2pll}$ for lower layer/subsurface flow.

4.4 Seasonal analysis

F_{logNSE} values separated for summer and winter differed from each other and also from those derived for the whole year (Fig. 4.7). Model performance was generally lower for the summer season, with $F_{logNSE} < 0.4$ for all models. Eight models had F_{logNSE} values below zero. They all used the same lower layer, subsurface flow and percolation structure combination as those model that performed poorest for the whole year. The models showing the best performance of summer recessions used all combinations including the TOPMODEL surface runoff structure S_{tmdl} (Fig. 4.8). However, in combination with $L_{unlimpow}$ for subsurface flow and lower layer models using S_{tmdl} performed poorer. The direct comparison of the performance for summer and winter resulted in a higher F_{logNSE} value for winter for almost all models. Two models showed the opposite. Both consist of a tension storage in the upper layer (either $U_{tension1}$ or $U_{tension2}$) and had exactly the same lower layer, subsurface flow/percolation structure ($L_{unlimfrc}$). All models where summer shows a better performance than winter use the percolation structure P_{f2sat}. All but one of the seven models with a F_{logNSE} less than zero in winter, used the percolation option P_{lower} in combination with $L_{unlimpow}$ for lower layer and subsurface flow. The same subsurface flow and lower layer option in combination with either P_{f2sat} or P_{w2sat} improved the model performance. Models using P_{lower} in combination with $L_{fixedsiz}$ had a high F_{logNSE}, and an even higher F_{logNSE} when S_{tmdl} was the surface runoff modeling option. The $L_{tens2pll}$ combined with any model option for the other structures always

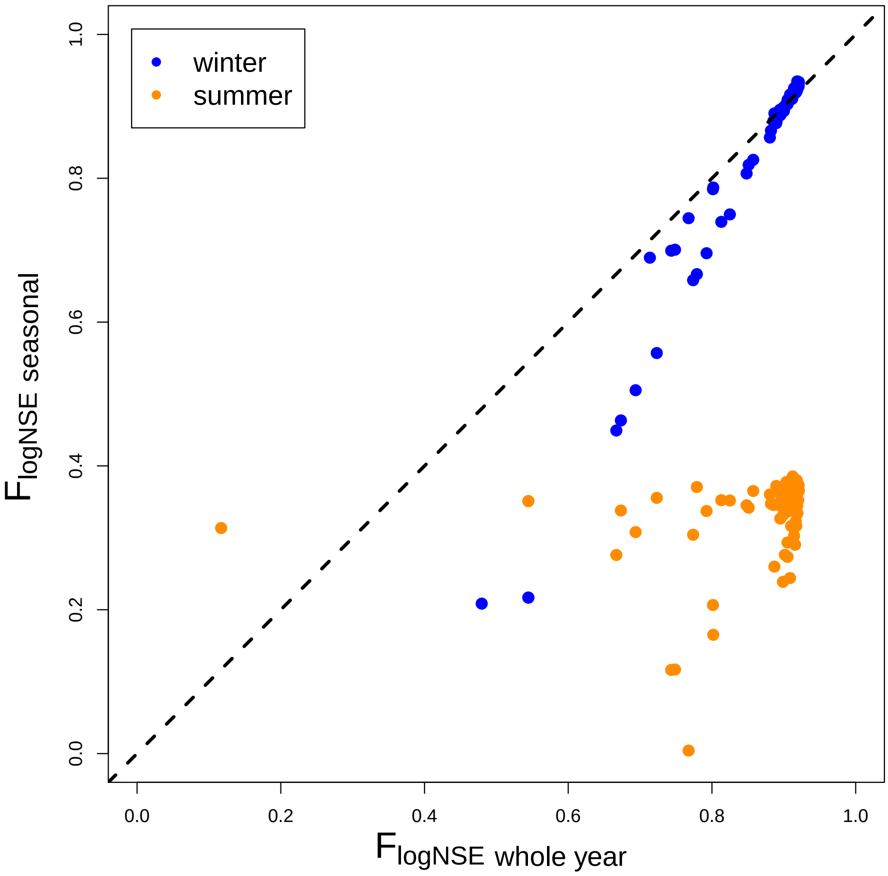

Figure 4.7: F_{logNSE} for summer and winter compared to F_{logNSE} for the whole year; the 8 models with $F_{logNSE} < 0$ are not shown.

performed better than a F_{logNSE} of 0.9 in winter. Most combinations of $L_{unlimfrc}$ with P_{f2sat} were found to range between F_{logNSE} 0.2 and 0.7. Combined with the surface runoff option S_{tmdl} it resulted in F_{logNSE} values of about 0.9. Generally, in summer observed recession slopes were steeper and flows were higher as compared to winter recessions which were slower with less steep slopes. Sometimes, a distinct non-linearity in recession slope was found with a considerably steeper recession slope from flow values of about 0.001 mm day^{-2} upwards. The recession relationships could be modelled with the polynomial (passed Kolmogorov-Smirnov-test) for 29 models for the winter season, for 44 for the whole year and for 28 models for the summer season. The polynomial described different recession relationships for summer and winter. The winter b and c coefficients of the polynomials are similar to those of the whole year. The structures of the underlying FUSE models were similar to the ones found for the whole year, but the lower layer and subsurface flow parametrisation were dominated by $L_{tens2pll_2}$. Only some models used $L_{unlimfrc}$, which was the dominant option for lower layer/subsurface flow for the whole year. In summer, more models had positive c coefficients and indeed there were cases where both coefficients were negative.

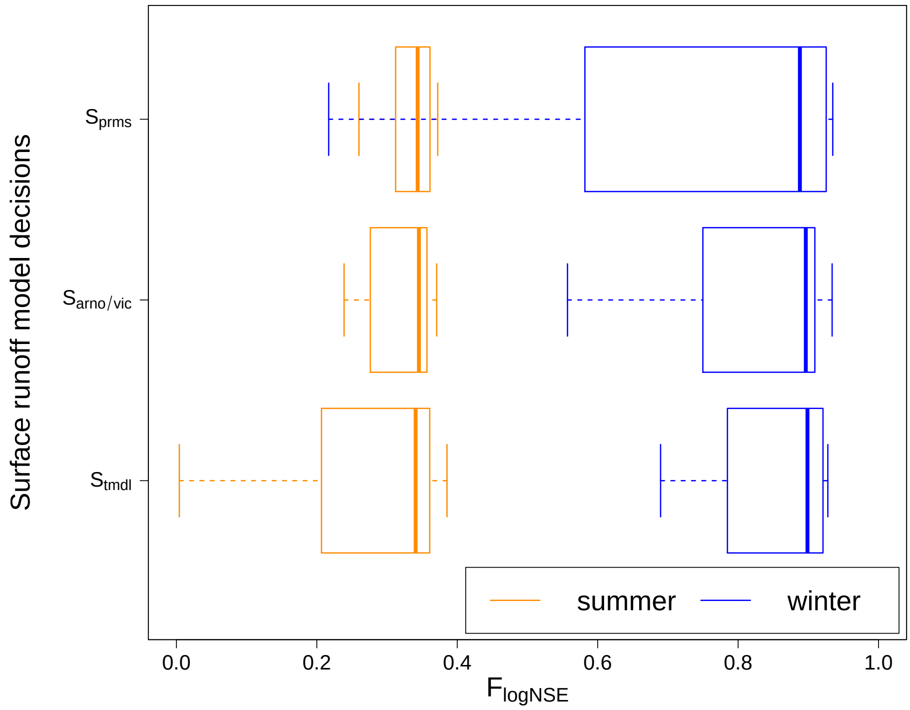

Figure 4.8: Boxplots of model performance for summer and winter streamflow simulations for the three surface runoff decision options.

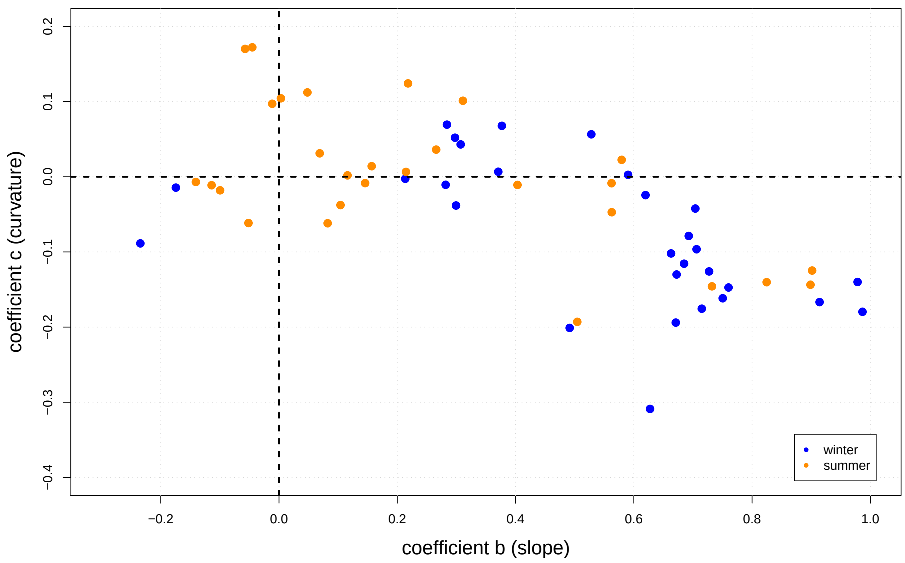

Figure 4.9: Coefficients of the polynomial fitted to seasonal $-dQ/dt$ to Q relationships.

5 Discussion

5.1 Model structures

The basic assumption in this study was that different model structures are the reason for the differences in model performance. Only four models performed well regarding the F_{logNSE} for both the whole year and for summer and winter. All used a combination of the lower layer/subsurface flow $L_{fixedsiz}$, upper layer $U_{tension2}$ and the percolation P_{f2sat}, containing at least two of the three components. For all other well performing models a systematic influence of a specific structural decision could not clearly be found. The models performed either better in one of the seasons or for the whole year.
Structural decisions that cause poor performance could be tracked based on the performance criteria F_{logNSE} and the simulation of the recession relationships. Such a structural decision is the lower layer/subsurface flow $L_{unlimpow}$ in combination with the percolation P_{lower}. This combination caused poor low flow simulations for the whole year as well as for the seasonal time series. Most of the binned versions of this combination could not be estimated using the polynomial as they did not pass the Kolomogorov Smirnov test. However, those that did pass, distinguished themselves by steep recession slopes.
The comparison of the slopes of summer and winter recessions reveals no seasonal differences for models with exactly this lower layer/subsurface flow and percolation combination. Clark et al. (2008) explain that here the lower layer is defined as a single state variable with no evaporation from this depth. The lower layer corresponds to the subsurface flow which is conceptualised by a power law originating from the parent model TOPMODEL. The main difference between the subsurface flow parametrisation in TOPMODEL and the other parent models is its dependency on the underlying distribution of the topographic index. The storage capacity in TOPMODEL also depends on the topographic index distribution and can hence be smaller or greater depending on the topography. In this study the Gamma distribution was used to define the distribution of the topographic index to keep some flexibility for calibration. Generally, the Gamma distribution is considered to be an appropriate assumption for the topographic index distribution of most catchments (Sivapalan et al., 1987). However, the models that used the TOPMODEL options may not have represented the topography in the Narsjø catchment well enough.
The percolation option P_{lower} is dependent on the lower layer decision. It thus strengthens the assumptions made with the lower layer/subsurface flow decision. The percolation option causes the fastest drainage when the lower layer is dry (Clark et al., 2008). Steep recession slopes were modelled with the combination of P_{lower} and $L_{unlimpow}$. The calibration with this combination appears to have caused a small water holding capacity of the lower layer resulting recessions that are steeper than recessions in the observed data. For the winter recessions of the models containing this combination for lower layer/ subsurface flow another fact should be kept in mind: in winter a snow storage is included. The precipitation data was pre-processed with the same snow routine for all FUSE models. Models input in winter is precipitation plus snow melt. Towards the end of the winter season (May/June) this process might fill the storages with small amounts of melt water and produce a prolongation of the recessions. The recessions modelled with

the combination of lower layer/subsurface flow and percolation options $L_{unlimpow}$ and P_{lower} are too fast and this results in unrealistic shapes of the recession relationships. The percolation option P_{lower} hence seems inappropriate for a combination with the lower layer/subsurface flow $L_{unlimpow}$ as it results in recessions that are too fast in summer and in streamflow that are too low in winter. None of the model decision combinations has such a distinct influence on model performance as the combination of $L_{unlimpow}$ and P_{lower}.

There are further combinations that systematically influence the seasonal performance: models containing the combination of $L_{unlimfrc}$ for lower layer and P_{f2sat} for percolation perform poorly for winter low flows. P_{f2sat} seems to influence the models ability to simulate low flows as it was used by all poorest performing models for winter. This means that the assumption of a percolation based on the field capacity should not be used to simulate winter recessions.

In summer, however, other model decisions cause a poor performance: one example is S_{tmdl} that models poor summer recessions. S_{tmdl} differs from other structures by surface runoff based on the distribution of the topographic index. Many model combinations in summer perform poorer when they contain the S_{tmdl} surface runoff. In summer, surface runoff plays a larger role for recessions than in winter.

Generally, model performance for low flows is easier to analyse for winter than for summer. In summer, there are several fast responding storages that contribute to the streamflow. The longer the recessions last, the less important become quickly draining storages that are prone to evaporation while slowly draining storages gain more influence. In addition, there can be a considerable influence by transpirating vegetation (Federer, 1973). In winter, the only storages that are important are lower layer storages and snow. Since only one snow storage option was modelled, only the lower layer storages matter. The results point out that the most important features for winter recession are directly connected to the lower storages. Hence, it is rather surprising to find a distinct modeling decision that causes a similar performance for both winter and summer recessions ($L_{unlimpow}$ plus P_{lower}).

In this study the choice of model structures was constrained to the structures of only four parent models. To keep the analysis manageable, in addition some processes were explicitly exempt, similar to the approach used in the original FUSE model (Clark et al., 2008). This includes climate input and hence required the preprocessing of the input data was with a snow accumulation and melt model instead of including several structural decisions of a snow model. Snow is in fact important in the Narsjø catchment, and testing structures describing the processes connected to snow might be worthwhile. This study, however, focused on the impact of model structures used to represent groundwater storage and release behavior. Future applications should consider testing more structures describing processes of snow melt and accumulation, but also interception and evapotranspiration, all of which were described with a single structural decision in this study. Further, the storage structural decisions included in this study are not the only options. Combinations of linear and non-linear reservoirs in series or parallel as tested in other studies could be appropriate for the Narsjø as well (e.g. Wang, 2011). Generally, it should be considered that an exclusion of alternative process representations using multiple hypothesis methods as FUSE can lead to the realization and evaluation of a model being biased by the modelers view (Clark et al., 2011).

5.2 Data quality

During the analysis some data issues common to winter streamflow measurements emerged. When ice forms in the river and at the gauging station, backwater effects may result due to ice blocking the channel. This will affect the validity of the rating curve or stop measuring devices alltogether requiring data gaps to be filled later (see e.g.(Moore et al., 2002)). A few mostly horizontal stripes can be seen in the Narjo data when plotting flow on a log scale (Fig. 4.5). However, here no gaps were filled (NVE, personal communication, 2010). Rupp and Selker (2006) also mention that measurement accuracy and changing rating curves in general may be the source of stripe-like patterns as in Fig. 4.5a. The difficulties of measuring low flows, particularly in winter, are well known and difficult to avoid. More detailed discussions can be found, for example, in Tallaksen and van Lanen (2004).
In general, validation of models with observed data of poor quality may lead to the rejection of models that might in fact be appropriate. A way to avoid the evaluation of model performance by standard metrics, such as the mean squared error, diagnostic signatures can be used (Yilmaz et al., 2008; Gupta et al., 2008). To include additional data on individual processes within a catchment may be necessary to identify scientifically defensible modeling strategies. Examples of application of diagnostic signatures in recession analysis can be found in e.g. McMillan et al. (2011) and Clark et al. (2011).

6 Conclusions

This study the impact of model structure on low flow simulations and recession behaviour has been assessed using the Framework for Understanding Structural Errors (FUSE). Using specific model structure combinations of different conceptual models resulted in different model performance for summer and winter low flows. Overall, individual structural decisions never appeared to be an exclusive reason, but rather the combinations of specific structural decisions affected model performance. Evaluating with F_{logNSE} as objective function, led to only a small number of models that performed well. While most well performing models did not allow for the detection of a systematic influence of a model structure combination on the model performance, poor performance was more clearly linked to specific model structures.
A specific structural combination for lower layer, subsurface flow and percolation was found that performed poorly in both seasons. The lower layer and subsurface flow structures influenced the winter low flow simulation, particularly. One main finding of this study was that there is a difference in model performance for summer and winter low flow and recession. In fact, all the structural decision combinations that were salient in this study were season specific – beside one combination that led to the poorest performance, independent on time period.
An important task would be to test this further for additional catchments with a seasonal flow regime (with snow in winter). In order to elucidate to which extent the influence of the considered model on low flow simulations are catchment specific or can be generalized, it should be replicated in other catchments. Those catchments should ideally be located in different topographical, geological and climatological regions.
The method itself, i.e. a systematic analysis of the structures of hydrological models

within the FUSE framework, using objective functions targeting at low flow and recession behaviour, seems promising. For low flow modelling it seems appropriate to use multiple objective functions and not to rely too much on a single function that is based on a comparison between simulated and observed data. Then, using FUSE allows to look at the model structures separately and to investigate the influence of the model structure on the model performance during low flow.

APPENDIX 5

Author contributions

A drought index accounting for snow (Appendix 3)
Maria Staudinger, Jan Seibert, and Kerstin Stahl
The basic idea to this research was born a discussion between all authors. Maria Staudinger developed and conceptualized this idea, did computations and analysis and wrote the article. Kerstin Stahl helped to revise the article and in particular with the restructuring of the discussion. Jan Seibert gave suggestions particularly for the elevation discretization of the snow model and helped to revise the article.

Quantifying sensitivity to droughts - an experimental modeling approach (Appendix 1)
Maria Staudinger, Markus Weiler, and Jan Seibert
The basic idea to this research developed in a discussion between all authors. Maria Staudinger developed and conceptualized this idea, did computations and analysis and wrote the article. Markus Weiler suggested the method to correct for precipitation in the scenarios and helped to revise the article. Jan Seibert helped to revise the article.

Comparison of hydrological model structures based on recession and low flow simulations (Appendix 4)
Maria Staudinger, Martyn Clark, Kerstin Stahl, Lena Tallaksen, and Jan Seibert
Maria Staudinger set up the concept of the study, did computations and analysis and wrote the article. Martyn Clark supervised the use of FUSE. Kerstin Stahl supervised the analysis. Jan Seibert discussed the results. All co-authors helped to revise the article.

Predictability of low flow - an assessment with simulation experiments (Appendix 2)
Maria Staudinger and Jan Seibert
Maria Staudinger developed and expanded the original idea, did computations and analysis and wrote the article. Jan Seibert had the incentive idea for this study, supervised and helped to revise the article.

References

Ahnert, F. (1987). “Approaches to dynamic equilibrium in theoretical simulations of slope development”. *Earth Surface Processes and Landforms* 12.1, pp. 3–15.

Ambroise, B., K. Beven, and J. Freer (1996). “Toward a generalization of the TOPMODEL concepts: Topographic indices of hydrological similarity”. *Water Resources Research* 32.7, pp. 2135–2145.

BAFU (2011). *Trockenheit Frühling 2011: Auswirkungen der Trockenheit auf Natur und Umwelt.* `http://www.bafu.admin.ch/dokumentation/fokus/11095/11332/index.html?lang=de`, last accessed 07.07.2014.

BUWAL, BWG, and MeteoSchweiz (2004). *Auswirkungen des Hitzesommers 2003 auf die Gewässer*. Schriftenreihe Umwelt Bundesamt für Umwelt, Wald und Landschaft, Bern.

Barnes, P. and P. Kalita (2001). “Watershed monitoring to address contamination source issues and remediation of the contaminant impairments”. *Water Science & Technology* 44.7, pp. 51–56.

Bathurst, J. C. et al. (2004). “Validation of catchment models for predicting land-use and climate change impacts. 3. Blind validation for internal and outlet responses”. *Journal of Hydrology* 287.1, pp. 74–94.

Beldring, S. et al. (2003). “Estimation of parameters in a distributed precipitation-runoff model for Norway”. *Hydrology and Earth System Sciences* 7.3, pp. 304–316.

Bergström, S. (1995). “The HBV model.” *Computer models of watershed hydrology.* Ed. by V.P. Singh, pp. 443–476.

Bergström, S. (1992). *The HBV model: Its structure and applications.* Swedish Meteorological and Hydrological Institute, p. 35.

Bergström, S., J. Harlin, and G. Lindström (1992). “Spillway design floods in Sweden: I. New guidelines”. *Hydrological Sciences Journal* 37.5, pp. 505–519.

Beven, K. (1984). “Infiltration into a class of vertically non-uniform soils”. *Hydrological Sciences Journal* 29.4, pp. 425–434.

— (1989). “Changing ideas in hydrology - the case of physically-based models”. *Journal of Hydrology* 105.1, pp. 157–172.

— (2001). *Rainfall-runoff modelling: the primer.* Wiley & Sons, New York, p. 372.

— (2006). “Searching for the Holy Grail of scientific hydrology:< i> Q< sub> t</sub></i>=(< i> S, R,? t</i>)< i> A</i> as closure”. *Hydrology and Earth System Sciences* 10.5, pp. 609–618.

— (2011). *Rainfall-runoff modelling: the primer, 2nd edition.* John Wiley & Sons, New York, p. 488.

Beven, K. and M. J. Kirkby (1979). “A physically based, variable contributing area model of basin hydrology”. *Hydrological Sciences Bulletin* 24.1, pp. 43–69.

Bierkens, M. F. P. and L. P. H. Van Beek (2009). “Seasonal predictability of European discharge: NAO and hydrological response time”. *Journal of Hydrometeorology* 10.4, pp. 953–968.

Boone, A. and P. J. Wetzel (1996). “Issues related to low resolution modeling of soil moisture: Experience with the PLACE model”. *Global and Planetary Change* 13.1-4, pp. 161–181.

Boulton, A. J. (2003). “Parallels and contrasts in the effects of drought on stream macroinvertebrate assemblages”. *Freshwater Biology* 48.7, pp. 1173–1185.

Box, G. E. and D. R. Cox (1964). “An analysis of transformed data”. *Journal of Royal Statistical Society. B* 39, pp. 211–252.

Bradford, M. J. and J. S. Heinonen (2008). “Low flows, instream flow needs and fish ecology in small streams”. *Canadian water resources Journal* 33.2, pp. 165–180.

Breuer, L. et al. (2009). “Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: Model intercomparison with current land use”. *Advances in Water Resources* 32.2, pp. 129–146.

Brutsaert, W. and J. L. Nieber (1977). “Regionalized drought flow hydrographs from a mature glaciated plateau”. *Water Resour. Res* 13.3, pp. 637–643.

Bunn, S. E. and A. H. Arthington (2002). “Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity”. *Environmental management* 30.4, pp. 492–507.

Bürgi, T. (2002). “Operational flood forecasting in mountainous areas-an interdisciplinary challenge”. *International Conference in Flood Estimation. CHR Report II-17. Bern, Switzerland*, pp. 397–406.

Burke, E. J. and S. J. Brown (2008). “Evaluating uncertainties in the projection of future drought”. *Journal of Hydrometeorology* 9.2, pp. 292–299.

Burke, E. J., S. J. Brown, and N. Christidis (2006). “Modeling the recent evolution of global drought and projections for the twenty-first century with the Hadley Centre climate model”. *Journal of Hydrometeorology* 7.5, pp. 1113–1125.

Burn, D. H. et al. (2008). “The processes, patterns and impacts of low flows across Canada”. *Canadian Water Resources Journal* 33.2, pp. 107–124.

Burnash, R. J. C. (1995). “The NWS river forecast system-catchment modeling”. *Computer Models of Watershed Hydrology*, pp. 311–366.

Cancelliere, A. and J. D. Salas (2010). “Drought probabilities and return period for annual streamflows series”. *Journal of Hydrology* 391.1, pp. 77–89.

Choi, H. T. and K. Beven (2007). “Multi-period and multi-criteria model conditioning to reduce prediction uncertainty in an application of TOPMODEL within the GLUE framework”. *Journal of Hydrology* 332.3, pp. 316–336.

Clark, M. P. and D. Kavetski (2010). “Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of time stepping schemes”. *Water Resources Research* 46.10, pp. 1–27.

Clark, M. P. et al. (2008). “Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models”. *Water Resources Research* 44.12. DOI: 10.1029/2007WR006735.

Clark, M. P. et al. (2009). "Consistency between hydrological models and field observations: linking processes at the hillslope scale to hydrological responses at the watershed scale". *Hydrological Processes* 23.2, pp. 311–319.

Clark, M. P. et al. (2011). "Hydrological field data from a modeller's perspective: Part 2: process-based evaluation of model hypotheses". *Hydrological Processes*, pp. 523–543.

Clausen, B. and B. Biggs (2000). "Flow variables for ecological studies in temperate streams: groupings based on covariance". *Journal of Hydrology* 237.3, pp. 184–197.

Cloke, H. L. and F. Pappenberger (2009). "Ensemble flood forecasting: a review". *Journal of Hydrology* 375.3, pp. 613–626.

Dawson, C. W., R. J. Abrahart, and L. M. See (2007). "HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts". *Environmental Modelling and Software* 22, pp. 1034–1052.

Demaria, E. M., B. Nijssen, and T. Wagener (2007). "Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model". *Journal of Geophysical Research: Atmospheres (1984–2012)* 112.D11. DOI: 10.1029/2006JD007534.

Dijk, A. I. J. M. et al. (2013). "The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society". *Water Resources Research*, pp. 1040–1057.

Doesken, N. J., T. B. McKee, and J. Kleist (1991). *Development of a surface water supply index for the western United States: Final report.* Colorado State University, Department of Atmospheric Science, pp. 164–175.

Dooge, J. C. I. (1992). "Sensitivity of runoff to climate change: A Hortonian approach". *Bulletin of the American Meteorological Society;(United States)* 73.12.

Duan, Q., S. Sorooshian, and V. K. Gupta (1994). "Optimal use of the SCE-UA global optimization method for calibrating watershed models". *Journal of Hydrology(Amsterdam)* 158.3, pp. 265–284.

Duan, Q. et al. (2006). "Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops". *Journal of Hydrology* 320.1-2, pp. 3–17.

Eltahir, E. A. B. and P. J-F. Yeh (1999). "On the asymmetric response of aquifer water level to floods and droughts in Illinois". *Water Resources Research* 35.4, pp. 1199–1217.

Engeland, K. (2002). "ECOMAG - Application to the Upper Glomma catchment (Four separate reports)". *Department of Geosciences, University of Oslo, Norway.*

Esko, K. (1980). "On the values and variability of degree-day melting factor in Finland". *Nordic hydrology* 11.5, pp. 235–242.

FOEN (2011). *Federal Office for the Environment, section hydrology.* accessed 20.01.2011. URL: http://www.bafu.admin.ch/hydrologie/12385/index.html?lang=en.

FOEN (2012). *Federal Office for the Environment, section hydrology.* last accessed: 12.12.2012. URL: http://www.bafu.admin.ch/hydrologie/12385/index.html?lang=en.

FOEN (2013a). *Federal Office for the Environment, section of hydrology.* http://www.bafu.admin.ch/hydrologie/12385/index.html?lang=en.

— (2013b). *Federal Office for the Environment, section of hydrology.* http://www.hydrodaten.admin.ch/en/. last accessed 20.12.2013.

Federer, C. A. (1973). “Forest transpiration greatly speeds streamflow recession”. *Water Resources Research* 9.6, pp. 1599–1604.

Feng, X., T. DelSole, and P. Houser (2011). “Bootstrap estimated seasonal potential predictability of global temperature and precipitation”. *Geophysical Research Letters* 38.L07702. DOI: 10.1029/2010GL046511.

Fleig, A. K. et al. (2006). “A global evaluation of streamflow drought characteristics”. *Hydrology and Earth System Sciences* 10.4, pp. 535–552.

Frei, C. (2013). “Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances”. *International Journal of Climatology*, pp. 1585–1605.

Fundel, F., S. Jörg-Hess, M. Zappa, et al. (2013). “Monthly hydrometeorological ensemble prediction of streamflow droughts and corresponding drought indices”. *Hydrol. Earth Syst. Sci* 17, pp. 395–407.

Ghosh, S. and P. P. Mujumdar (2007). “Nonparametric methods for modeling GCM and scenario uncertainty in drought assessment”. *Water Resources Research* 43.7, W07405. DOI: 10.1029/2006WR005351.

Gibbs, W. J. and J. V. Maher (1967). “Rainfall deciles as drought indicators”. *Bureau of Meteorology Bulletin* 48, p. 33.

Gobena, A. K. and T. Y. Gan (2010). “Incorporation of seasonal climate forecasts in the ensemble streamflow prediction system”. *Journal of Hydrology* 385.1-4, pp. 336–352.

Gottschalk, L., L. M. Tallaksen, and G. Perzyna (1997). “Derivation of low flow distribution functions using recession curves”. *Journal of Hydrology* 194.1-4, pp. 239–262.

Graf, W. L. (1977). “The rate law in fluvial geomorphology”. *American Journal of Science* 277.2, pp. 178–191.

Grijsen, J. G., X. C. Snoeker, and C. J. M. Vermeulen (1993). *An information system for flood early warning.* Delft Hydraulics Laboratory, pp. 263–289.

Gudmundsson, L. et al. (2012). “Comparing large-scale hydrological model simulations to observed runoff percentiles in Europe”. *Journal of Hydrometeorology* 13.2, pp. 604–620.

Gupta, H. V. et al. (2009). “Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling”. *Journal of Hydrology* 377.1, pp. 80–91.

Gupta, H., K. Beven, and T. Wagener (2005). “Model calibration and uncertainty estimation”. *Encyclopedia of Hydrological Sciences*, pp. 2012–2031.

Gupta, H.V., T. Wagener, and Y. Liu (2008). “Reconciling theory with observations: elements of a diagnostic approach to model evaluation”. *Hydrological Processes* 22.18, pp. 3802–3813.

Gustard, A., A. Bullock, and J. M. Dixon (1992). *Low flow estimation in the United Kingdom.* Institute of Hydrology.

Gustard, A. et al. (2008). *Manual on low flow estimation and prediction.* Tech. rep. 50. World Meteorological Organization, Geneva, Switzerland.

Guttman, N. B. (1999). “Accepting the Standardized Precipitation Index: a calculation algorithm”. *Journal of the American Water Resources Association* 35.2, pp. 311–322.

Hall, F. R. (1968). “Base-Flow RecessionsŮA Review”. *Water Resources Research* 4.5, pp. 973–983.

Harman, C. J., M. Sivapalan, and P. Kumar (2009). "Power law catchment-scale recessions arising from heterogeneous linear small-scale dynamics". *Water resources research* 45.9, W09404. DOI: 10.1029/2008WR007392.

Haslinger, K. et al. (2014). "Exploring the link between meteorological drought and streamflow: Effects of climate-catchment interaction". *Water Resources Research*, pp. 2468–2487.

Hayes, M. J. et al. (1999). "Monitoring the 1996 drought using the standardized precipitation index". *Bulletin of the American Meteorological Society* 80.3, pp. 429–438.

Heim Jr, R. R. (2002). "A review of twentieth-century drought indices used in the United States". *Bulletin of the American Meteorological Society* 83.8, pp. 1149–1165.

Henderson-Sellers, A., Z. L. Yang, and R. E. Dickinson (1993). "The project for intercomparison of land-surface parameterization schemes". *Bulletin of the American Meteorological Society* 74.7, pp. 1335–1349.

Hipel, K. W. and A. I. McLeod (1994). *Time series modelling of water resources and environmental systems.* 45. Elsevier Science Limited, p. 1012.

Hock, R. (2003). "Temperature index melt modelling in mountain areas". *Journal of Hydrology* 282.1, pp. 104–115.

Holländer, H. M. et al. (2009). "Comparative predictions of discharge from an artificial catchment(Chicken Creek) using sparse data". *Hydrology and Earth System Sciences* 13.11, pp. 2069–2094.

Hosking, J. R. M. (1990). "L-moments: analysis and estimation of distributions using linear combinations of order statistics". *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 105–124.

Institute of Hydrology (1980). *Low flow studies report.*

Ji, L. and A. J. Peters (2003). "Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices". *Remote Sensing of the Environment* 87.1, pp. 85–98.

Jordan, T. E., D. L. Correll, and D. E. Weller (1997). "Relating nutrient discharges from watersheds to land use and streamflow variability". *Water Resources Research* 33.11, pp. 2579–2590.

Kavetski, D. and G. Kuczera (2007). "Model smoothing strategies to remove microscale discontinuities and spurious secondary optima in objective functions in hydrological calibration". *Water Resources Research* 43.3, W03411. DOI: 10.1029/2006WR005195.

Keyantash, J. and J. A. Dracup (2002). "The quantification of drought: an evaluation of drought indices". *Bulletin of the American Meteorological Society* 83.8, pp. 1167–1180.

Kirchner, J. W. (2009). "Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward". *Water Resources Research* 45.2. DOI: 10.1029/2008WR006912.

Ko, M-K. and A. Tarhule (1994). "Streamflow droughts of northern Nigerian rivers". *Hydrological Sciences Journal* 39.1, pp. 19–34.

Kohn, I. et al. (2014). "Low flow in Germany in 2011". *Hydrologie und Wasserwirtschaft.*

Krause, P., D. P. Boyle, and F. Bäse (2005). "Comparison of different efficiency criteria for hydrological model assessment". *Advances in Geosciences* 5, pp. 89–97.

Kroll, C. et al. (2004). "Developing a watershed characteristics database to improve low streamflow prediction". *Journal of Hydrologic Engineering* 9.2, pp. 116–125.

Lamb, R., K. Beven, et al. (1997). "Using interactive recession curve analysis to specify a general catchment storage model". *Hydrology and Earth System Sciences* 1.1, pp. 101–113.

Leavesley, G. H. et al. (1983). "Precipitation-runoff modeling system users manual". *USGS Series Water-Resources Investigations Report*, p. 207.

Legates, D. R. and G. J. McCabe Jr (1999). "Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation". *Water Resources Research* 35.1, pp. 233–241.

Lehner, B. et al. (2006). "Estimating the impact of global change on flood and drought risks in Europe: a continental, integrated analysis". *Climatic Change* 75.3, pp. 273–299.

Lettenmaier, D. P. et al. (1999). "Water resources implications of global warming: A US regional perspective". *Climatic Change* 43.3, pp. 537–579.

Li, H. et al. (2009). "The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting". *Journal of Geophysical Research* 114.D4, p. D04114. DOI: 10.1029/2008JD010969.

Li, J. et al. (2010). "An extreme drought event in the central Tien Shan area in the year 1945". *Journal of Arid Environments* 74.10, pp. 1225–1231.

Liang, X. et al. (1994). "A simple hydrologically based model of land surface water and energy fluxes for general circulation models". *J. Geophys. Res* 99.14, pp. 415–14.

Lindström, G. (1997). "A simple automatic calibration routine for the HBV model". *Nordic Hydrology* 28.3, pp. 153–168.

Lindström, G. et al. (1997a). "Development and test of the distributed HBV-96 hydrological model". *Journal of Hydrology* 201.1, pp. 272–288.

— (1997b). "Development and test of the distributed HBV-96 hydrological model". *Journal of hydrology* 201.1, pp. 272–288.

López-Moreno, J. I. et al. (2009). "Downstream propagation of hydrological droughts in highly regulated transboundary rivers: the case of the Tagus River between Spain and Portugal". *Water Resources Research* 45, W02405.

Mahanama, S. et al. (2012). "Soil moisture, snow, and seasonal streamflow forecasts in the United States". *Journal of Hydrometeorology* 13.1, pp. 189–203.

Massey Jr, F.J. (1951). "The Kolmogorov-Smirnov test for goodness of fit". *Journal of the American Statistical Association* 46.253, pp. 68–78.

Matonse, A. H. and C. Kroll (2009). "Simulating low streamflows with hillslope storage models". *Water Resources Research* 45.1, W01407. DOI: 10.1029/2007WR006529.

McGlynn, B. L. et al. (2003). "On the relationships between catchment scale and streamwater mean residence time". *Hydrological Processes* 17.1, pp. 175–181.

McGuire, K. J. et al. (2005). "The role of topography on catchment-scale water residence time". *Water Resources Research* 41.5, W05002. DOI: 10.1029/2004WR003657.

McKee, T.B., N. J. Doesken, and J. Kleist (1993). "The relationship of drought frequency and duration to time scales". *Proceedings of the 8th Conference on Applied Climatology*. Vol. 17. American Meteorological Society Boston, MA, pp. 179–183.

McMahon, T. A. et al. (2012). "Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis". *Hydrology and Earth System Sciences* 9.10, pp. 11829–11910.

McMillan, H. K. et al. (2011). “Hydrological field data from a modeller’s perspective: Part 1. Diagnostic tests for model structure”. *Hydrological Processes*, pp. 511–522.

Melching, C. S. (1995). “Reliability estimation”. *Computer models of watershed hydrology*, pp. 69–118.

Merz, R. and G. Blöschl (2004). “Regionalisation of catchment model parameters”. *Journal of Hydrology* 287.1, pp. 95–123.

MeteoSwiss (2013). *Federal Office of Meteorology and Climatology.* `http://www.meteoschweiz.admin.ch/web/en/services/data_portal/gridded_datasets.html`. last accessed 10.10.2013.

MeteoSwiss (2013a). *Federal Office of Meteorology and Climatology.* `http://www.meteoschweiz.admin.ch/web/de/services/datenportal/gitterdaten/precip/rhiresd.html`. last accessed: 10.07.2013.

— (2013b). *Federal Office of Meteorology and Climatology.* `http://www.meteoschweiz.admin.ch/web/de/services/datenportal/gitterdaten/temp/tabsd.html`. last accessed: 10.07.2013.

Miller, N. L., K. E. Bashford, and E. Strem (2003). “Potential impacts of climate change on California Hydrology”. *JAWRA Journal of the American Water Resources Association* 39.4, pp. 771–784.

Mishra, A. K. and V. P. Singh (2010). “A review of drought concepts”. *Journal of Hydrology* 391.1, pp. 202–216.

Montanari, A. (2005). “Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations”. *Water Resources Research* 41.8.

Moore, D.S. and W.I. Notz (2005). *Statistics: Concepts and controversies.* WH Freeman & Co.

Moore, R. D., A. S. Hamilton, and J. Scibek (2002). “Winter streamflow variability, Yukon Territory, Canada”. *Hydrological Processes* 16.4, pp. 763–778.

Moore, R.J. and D.A. Jones (1997). “Linking hydrological and hydrodynamic forecast models and their data”. *RIBAMOD River Basin Modeling, management and flood mitigation: Proceedings of the first workshop.* European Community, EUR17456EN, pp. 37–56.

Naresh Kumar, M. et al. (2009). “On the use of Standardized Precipitation Index (SPI) for drought intensity assessment”. *Meteorological Applications* 16.3, pp. 381–389.

Nash, J.E. and J.V. Sutcliffe (1970). “River flow forecasting through conceptual models - part I- a discussion of principles”. *Journal of Hydrology* 10, pp. 282–290.

Nijssen, B. et al. (2001). “Hydrologic sensitivity of global rivers to climate change”. *Climatic change* 50.1-2, pp. 143–175.

Null, S. E., J. H. Viers, and J. F. Mount (2010). “Hydrologic response and watershed sensitivity to climate warming in California’s Sierra Nevada”. *PLoS One* 5.4, e9932. DOI: `10.1371/journal.pone.0009932`.

Orlowsky, B. and S. I. Seneviratne (2012). “Global changes in extreme events: regional and seasonal dimension”. *Climatic Change* 110.3, pp. 669–696.

Pagano, T. and D. Garen (2005). “A recent increase in western US streamflow variability and persistence”. *Journal of Hydrometeorology* 6.2, pp. 173–179.

Palmer, W. C. (1965). *Meteorological drought.* US Department of Commerce, Weather Bureau Washington DC, USA.

Palmroth, S., G.G. Katul, and R. Oren (2010). “Estimation of long-term basin scale evapotranspiration from streamflow time series”. *Water Resources Research* 46, pp. 1–13.

Pappenberger, F. and K. Beven (2004). “Functional classification and evaluation of hydrographs based on Multicomponent Mapping (Mx)”. *International Journal of River Basin Management* 2.2, pp. 89–100.

Parkin, G. et al. (1996). “Validation of catchment models for predicting land-use and climate change impacts. 2. Case study for a Mediterranean catchment”. *Journal of Hydrology* 175.1, pp. 595–613.

Pearson, K. (1896). “Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia”. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, pp. 253–318.

Penman, H. L. (1948). “Natural evaporation from open water, bare soil and grass”. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 193.1032, pp. 120–145.

Perrin, C., C. Michel, and V. Andréassian (2001). “Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments”. *Journal of Hydrology* 242.3-4, pp. 275–301.

Peters, E. et al. (2003). “Propagation of drought through groundwaterŮa new approach using linear reservoir theory”. *Hydrological processes* 17.15, pp. 3023–3040.

Press, W.H. et al. (1992). *Numerical recipes in FORTRAN: the art of scientific computing.* Cambridge Univ Press.

Price, K. (2011). “Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions: A review”. *Progress in Physical Geography* 35.4, pp. 465–492.

Pushpalatha, R. et al. (2012). “A review of efficiency criteria suitable for evaluating low-flow simulations”. *Journal of Hydrology* 420, pp. 171–182.

Pyrce, R. (2004). “Hydrological low flow indices and their uses”. *Watershed Science Centre,(WSC) Report* 04-2004, p. 33.

Rango, A. and J. Martinec (1995). “Revisiting the degree-day method for snowmelt computations”. *JAWRA Journal of the American Water Resources Association* 31.4, pp. 657–669.

Rasmussen, R. et al. (2012). “How Well Are We Measuring Snow: The NOAA/FAA/NCAR Winter Precipitation Test Bed.” *Bulletin of the American Meteorological Society* 93.6.

Rebetez, M. et al. (2006). “Heat and drought 2003 in Europe: a climate synthesis”. *Annals of Forest Science* 63.6, pp. 569–577.

Reed, S. et al. (2004). “Overall distributed model intercomparison project results”. *Journal of Hydrology* 298.1, pp. 27–60.

Reusser, D. E. et al. (2009). “Analysing the temporal dynamics of model performance for hydrological models”. *Hydrol. Earth Syst. Sci* 13, pp. 999–1018.

Roering, J. J., J. W. Kirchner, and W. E. Dietrich (2001). “Hillslope evolution by nonlinear, slope-dependent transport: Steady state morphology and equilibrium adjustment timescales”. *Journal of Geophysical Research: Solid Earth (1978–2012)* 106.B8, pp. 16499–16513.

Rupp, D. E. and J.S. Selker (2006). “Information, artifacts, and noise in dQ/dt- Q recession analysis”. *Advances in Water Resources* 29.2, pp. 154–160.

Rupp, D. E. and Ross A Woods (2008). “Increased flexibility in base flow modelling using a power law transmissivity profile”. *Hydrological Processes* 22.14, pp. 2667–2671.

Sankarasubramanian, A., R. M. Vogel, and J. F. Limbrunner (2001). “Climate elasticity of streamflow in the United States”. *Water Resources Research* 37.6, pp. 1771–1781.

Santos, J., J. Corte-real, and S. Leite (2007). “Atmospheric large-scale dynamics during the 2004/2005 winter drought in portugal”. *International journal of climatology* 27.5, pp. 571–586.

Sawicz, K. et al. (2011). “Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA”. *Hydrology and Earth System Sciences* 8.3, pp. 4495–4534.

Schaake, J. C. and P.E. Waggoner (1990). “From climate to flow.” *Climate change and US water resources.* Pp. 177–206.

Schaefli, B. and H. V. Gupta (2007). “Do Nash values have value?” *Hydrological Processes* 21.15, pp. 2075–2080.

Schär, C. et al. (2004). “The role of increasing temperature variability in European summer heatwaves”. *Nature* 427.6972, pp. 332–336.

Schlegel, T., H. Kunz, and D. Gerstgrasser (2011). *Aktuelles zum Wettergeschehen - Trockenheit Jahresbeginn 2011.* `http://www.meteoschweiz.admin.ch/web/de/wetter/wetterereignisse/trockenheit_jahresbeginn.html`, last accessed 07.07.2014.

Schorer, M. (1992). *Extreme Trockensommer in der Schweiz und ihre Folgen für Natur und Wirtschaft.* Dissertation, Geographisches Institut der Universität Bern.

Schüpp, M. and G. Gensler (2012). *Klimaregionen der Schweiz, in: Die Beobachtungsnetze der Schweizerischen Meteorologischen Anstalt, Konzept 1980.* Arbeitsberichte der Schweizerischen Meteorologischen Anstalt, Anhang Ib. Zürich.

Seibert, J. (1999). “Regionalisation of parameters for a conceptual rainfall-runoff model”. *Agricultural and Forest Meteorology* 98.1, pp. 279–293.

— (2000). “Multi-criteria calibration of a conceptual runoff model using a genetic algorithm”. *Hydrology and Earth System Sciences* 4.2, pp. 215–224.

— (2001). “On the need for benchmarks in hydrological modelling”. *Hydrological Processes* 15.6, pp. 1063–1064.

Seibert, J. and J. J. McDonnell (2010). “Land-cover impacts on streamflow: a change-detection modelling approach that incorporates parameter uncertainty”. *Hydrological Sciences Journal–Journal des Sciences Hydrologiques* 55, pp. 316–332.

Seibert, J. and M. J. P. Vis (2012). “Teaching hydrological modeling with a user-friendly catchment-runoff-model software package”. *Hydrol. Earth Syst. Sci* 16, pp. 3315–3325.

Seneviratne, S. I. et al. (2013). “Trockenheit in der Schweiz”. *Aqua & Gas* 9, pp. 38–47.

Shafer, B. A. and L. E. Dezman (1982). “Development of a Surface Water Supply Index (SWSI) to assess the severity of drought conditions in snowpack runoff areas”. *Proceedings of the Western Snow Conference.* Vol. 50, pp. 164–175.

Sheffield, Justin and Eric F Wood (2012). *Drought: past problems and future scenarios.* Routledge, p. 192.

Shukla, S. and D. P. Lettenmaier (2011). "Seasonal hydrologic prediction in the United States: understanding the role of initial hydrologic conditions and seasonal climate forecast skill". *Hydrology and Earth System Sciences* 15.11, pp. 6565–6592.

Shukla, S. and A. W. Wood (2008). "Use of a standardized runoff index for characterizing hydrologic drought". *Geophysical Research Letters* 35.2, p. L02405. DOI: 10.1029/2007GL032487.

Singla, S. et al. (2012). "Predictability of soil moisture and river flows over France for the spring season". *Hydrol. Earth Syst. Sci* 16, pp. 201–216.

Sivapalan, M., K. Beven, and E. F. Wood (1987). "On hydrologic similarity 2. A scaled model of storm runoff production". *Water Resources Research* 23.12, pp. 2266–2278.

Smakhtin, V. U. (2001). "Low flow hydrology: a review". *Journal of hydrology* 240.3, pp. 147–186.

Solomatine, D.P. and T. Wagener (2011). "2.16 - Hydrological Modeling". *Treatise on Water Science*. Ed. by P. Wilderer. Oxford: Elsevier, pp. 435 –457. ISBN: 978-0-444-53199-5. DOI: http://dx.doi.org/10.1016/B978-0-444-53199-5.00044-0. URL: http://www.sciencedirect.com/science/article/pii/B9780444531995000440.

Solomon, S. (2007). *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC*. Vol. 4. Cambridge University Press.

Sorooshian, S. and V. K. Gupta (1995). *Model calibration. Computer Models of Watershed Hydrology*. Ed. by V. P. Singh. Water Resources Publications, pp. 23–63.

Spreafico, M., C. Leibundgut, and R. Weingartner (1992). *Hydrologischer Atlas der Schweiz: Atlas hydrologique de la Suisse*. Bundesamt für Landestopographie.

Stagge, J. H. et al. (2013). *A European Drought Reference (EDR) Database: design and Online Implementation*. Tech. rep. 12, p. 42. URL: www.eu-drought.org/technicalreports.

Stahl, K. and S. Demuth (1999). "Linking streamflow drought to the occurrence of atmospheric circulation patterns". *Hydrological sciences journal* 44.3, pp. 467–482.

Stahl, K. et al. (2012). *A European Drought Impact Report Inventory (EDII): Design and Test for Selected Recent Droughts in Europe DROUGHT-R&SPI*. Tech. rep. 3. URL: http://www.eu-drought.org/technicalreports.

Stoelzle, M., K. Stahl, and M. Weiler (2013). "Are streamflow recession characteristics really characteristic?" *Hydrology and Earth System Sciences* 17.2, pp. 817–828.

Stoelzle, M. et al. (2014a). "Is there a superior conceptual groundwater model structure for baseflow simulation?" *Hydrological Processes*. DOI: 10.1002/hyp.10251.

Stoelzle, M. et al. (2014b). "Streamflow sensitivity to drought scenarios in catchments with different geology". *Geophysical Research Letters*. DOI: 10.1002/2014GL061344.

Sujono, J., S. Shikasho, and K. Hiramatsu (2004). "A comparison of techniques for hydrograph recession analysis". *Hydrological Processes* 18, pp. 403–413.

Szilagyi, J., Z. Gribovszki, and P. Kalicz (2007). "Estimation of catchment-scale evapotranspiration from baseflow recession data: Numerical model and practical application results". *Journal of Hydrology* 336.1, pp. 206–217.

Szilagyi, Jozsef, Marc B Parlange, and John D Albertson (1998). "Recession flow analysis for aquifer parameter determination". *Water Resources Research* 34.7, pp. 1851–1857.

Tague, C. and G. E. Grant (2009). “Groundwater dynamics mediate low-flow response to global warming in snow-dominated alpine regions”. *Water Resources Research* 45.7, W07421. DOI: 10.1029/2008WR007179.

Tallaksen, L. M. (1995). “A review of baseflow recession analysis”. *Journal of Hydrology* 165, pp. 349–370.

Tallaksen, L. M. and H. A. J. van Lanen (2004). *Hydrological drought: processes and estimation methods for streamflow and groundwater*. Elsevier, p. 579.

Tate, E. L. and S. N. Freeman (2000). “Three modelling approaches for seasonal streamflow droughts in southern Africa: the use of censored data”. *Hydrological Sciences Journal* 45.1, pp. 27–42.

Teegavarapu, R. and A. Elshorbagy (2005). “Fuzzy set based error measure for hydrologic model evaluation”. *Journal of Hydroinformatics* 7, pp. 199–203.

Tetzlaff, D., J. Seibert, and C. Soulsby (2009). “Inter-catchment comparison to assess the influence of topography and soils on catchment transit times in a geomorphic province; the Cairngorm mountains, Scotland”. *Hydrological Processes* 23.13, pp. 1874–1886.

Teutschbein, C. and J. Seibert (2012). “Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods”. *Journal of Hydrology* 456, pp. 12–29.

— (2013). “Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions?” *Hydrology and Earth System Sciences* 17.12, pp. 5061–5077.

Trigo, R. M., C. M. Gouveia, and D. Barriopedro (2010). “The intense 2007–2009 drought in the Fertile Crescent: Impacts and associated atmospheric circulation”. *Agricultural and Forest Meteorology* 150.9, pp. 1245–1257.

Uhlenbrook, S. et al. (1999). “Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure”. *Hydrological Sciences Journal* 44.5, pp. 779–797.

Van Kalken, T., C. Skotner, and H. Madsen (2004). “A new generation, GIS based, open flood forecasting system”. *Proceedings of the 8th National conference on hydraulics in Water Engineering, The institute of Engineers: Australia*, p. 8.

Van Loon, A. F. (2013). *On the propagation of drought: how climate and catchment characteristics influence hydrological drought development and recovery.* Dissertation, Wageningen University.

Van Loon, A. F. and H. A. J. Van Lanen (2011). “A process-based typology of hydrological drought”. *Hydrology and Earth System Sciences* 8, pp. 11413–11483.

Van Loon, A. F. et al. (2010). “Understanding hydrological winter drought in Europe”. *Global Change: Facing Risks and Threats to Water Resources, IAHS Publ* 340, pp. 189–197.

Van Loon, A. F. et al. (2014). “How climate seasonality modifies drought duration and deficit”. *Journal of Geophysical Research: Atmospheres* 119.8, pp. 4640–4656. DOI: 10.1002/2013JD020383. URL: http://dx.doi.org/10.1002/2013JD020383.

Van Loon, A.F. and H.A.J. Van Lanen (2012). “A process-based typology of hydrological drought”. *Hydrology and Earth System Sciences* 16.7, pp. 1915–1946.

Van Vliet, M. et al. (2012). “Vulnerability of US and European electricity supply to climate change”. *Nature Climate Change* 2.9, pp. 676–681.

Vicente-Serrano, S. M., S. Beguería, and J. I. López-Moreno (2010). "A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index". *Journal of Climate* 23.7, pp. 1696–1718.

Vicente-Serrano, S. M. et al. (2011). "Accurate computation of a streamflow drought index". *Journal of Hydrologic Engineering* 17.2, pp. 318–332.

Vicente-Serrano, S. M. et al. (2012). "Performance of Drought Indices for Ecological, Agricultural, and Hydrological Applications." *Earth Interactions* 16.10, pp. 1–27.

Vidal, J. P. et al. (2010). "Multilevel and multiscale drought reanalysis over France with the Safran-Isba-Modcou hydrometeorological suite". *Hydrology and Earth System Sciences* 14.3, pp. 459–478.

Viste, E., D. Korecha, and A. Sorteberg (2013). "Recent drought and precipitation tendencies in Ethiopia". *Theoretical and Applied Climatology*, pp. 1–17.

Viviroli, D. et al. (2009). "An introduction to the hydrological modelling system PREVAH and its pre-and post-processing-tools". *Environmental Modelling & Software* 24.10, pp. 1209–1222.

Vogel, R. M. and N. M. Fennessey (1994). "Flow-duration curves. I: New interpretation and confidence intervals". *Journal of Water Resources Planning and Management* 120.4, pp. 485–504.

Vogel, R. M., Y. Tsai, and J. F. Limbrunner (1998). "The regional persistence and variability of annual streamflow in the United States". *Water Resources Research* 34.12, pp. 3445–3459.

Vogel, R.M. and N.M. Fennessey (1995). "Flow Duration Curves II: A Review of Applications in Water Resources Planning". *Journal of the American Water Resources Association* 31.6, pp. 1029–1039.

WMO (1974). *International Glossary of Hydrology*. Geneva: World Meteorological Organization.

Wagener, T. and H. S. Wheater (2004). *Rainfall-runoff modelling in gauged and ungauged catchments*. Imperial College Press, p. 332.

— (2006). "Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty". *Journal of Hydrology* 320.1, pp. 132–154.

Wagener, T. et al. (2007). "Catchment classification and hydrologic similarity". *Geography Compass* 1.4, pp. 901–931.

Wang, D. (2011). "On the base flow recession at the Panola Mountain Research Watershed, Georgia, United States". *Water Resources Research* 47.3, pp. 1–10.

Wang, D. and X. Cai (2009). "Detecting human interferences to low flows through base flow recession analysis". *Water resources research* 45.7, W07426. DOI: 10.1029/2009WR007819.

Wang, G. (2005). "Agricultural drought in a future climate: results from 15 global climate models participating in the IPCC 4th assessment". *Climate dynamics* 25.7-8, pp. 739–753.

Westra, S. and A. Sharma (2010). "An upper limit to seasonal rainfall predictability?" *Journal of Climate* 23.12, pp. 3332–3351.

Wetherald, R. T. and S. Manabe (1999). "Detectability of summer dryness caused by greenhouse warming". *Climatic change* 43.3, pp. 495–511.

— (2002). "Simulation of hydrologic changes associated with global warming". *Journal of Geophysical Research* 107.D19, p. 4379. DOI: 10.1029/2001JD001195.

Wilby, R. L. and I. Harris (2006). "A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK". *Water Resources Research* 42.2, W02419. DOI: 10.1029/2005WR004065.

Wilhite, D. A. (2002). "Combating drought through preparedness". *Natural Resources Forum*. Vol. 26. 4. Wiley Online Library, pp. 275–285.

Wittenberg, H. (1999). "Baseflow recession and recharge as nonlinear storage processes". *Hydrological Processes* 13.5, pp. 715–726.

— (2003). "Effects of season and man-made changes on baseflow and flow recession: case studies". *Hydrological Processes* 17.11, pp. 2113–2123.

Wittenberg, H. and M. Sivapalan (1999). "Watershed groundwater balance estimation using streamflow recession analysis and baseflow separation". *Journal of Hydrology* 219.1, pp. 20–33.

Wood, A. W. and D. P. Lettenmaier (2008). "An ensemble approach for attribution of hydrologic prediction uncertainty". *Geophysical Research Letters* 35.14, p. L14401. DOI: 10.1029/2008GL034648.

Yilmaz, K. K., H. V. Gupta, and T. Wagener (2008). "A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model". *Water Resources Research* 44, W09417. DOI: 10.1029/2007WR006716.

Zaidman, M. D., H. G. Rees, A. R. Young, et al. (2002). "Spatio-temporal development of streamflow droughts in north-west Europe". *Hydrology and Earth System Sciences* 6.4, pp. 733–751.

Zelenhasic, E. and A. Salvai (1987). "Method of Streamflow Drought Analysis". *Water Resources Research* 23.1, pp. 156–168.

Zhao, RJ (1977). "Flood forecasting method for humid regions of China". *East China Institute of Hydraulic Engineering, Nanjing, China.*